

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

10/534538 534538

(43) 国际公布日:  
2004年6月24日(24.06.2004)

PCT

(10) 国际公布号:  
WO 2004/052910 A1(51) 国际分类号<sup>7</sup>: C07H 21/04, C07K 14/465, 14/78, C12N 15/63, A61K 38/17, 38/39, A61P 19/02OFFICE); 中国北京市阜成门外大街2号万通新世界  
广场8层, Beijing 100037 (CN).

(21) 国际申请号: PCT/CN2003/000967

(22) 国际申请日: 2003年11月14日(14.11.2003)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权: 02149375.8 2002年11月14日(14.11.2002) CN

(71) 申请人(对除美国以外的所有指定国): 中国人民解放军军事医学科学院附属医院(AFFILIATED HOSPITAL OF ACADEMY OF MILITARY MEDICAL SCIENCES, PLA) [CN/CN]; 中国北京市海淀区北太平路2号, Beijing 100039 (CN)。

(81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(84) 指定国(地区): ARIPO专利(BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

(72) 发明人: 及

(73) 发明人/申请人(仅对美国): 羿永志(XI, Yongzhi) [CN/CN]; 习彩霞(XI, Caixia) [CN/CN]; 中国北京市海淀区北太平路2号, Beijing 100039 (CN)。

(74) 代理人: 中国国际贸易促进委员会专利商标事务所  
(CCPIT PATENT AND TRADEMARK LAW)本国国际公布:  
— 包括国际检索报告。所引用双字母代码和其它缩写符号, 请参考刊登在每期  
PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: A FULL-LENGTH POLYNUCLEOTIDE CODING CHICKEN TYPE II COLLAGEN AND THE USE OF IT

(54) 发明名称: 一种编码鸡II型胶原蛋白的全长多核苷酸序列及其用途

(57) Abstract: The invention disclosed a nucleic acid molecule including a full-length polynucleotide coding chicken type II collagen which can be seen from SEQ ID NO:1, a segment having the same biological function with it; and a chicken type II collagen coded by it. The invention also disclosed a method of preparing this chicken type II collagen, and the use of this chicken type II collagen in preparing the pharmaceutical compound which can be used to treat or prevent RA. The invention specially involved a pharmaceutical compound which can be used to treat or prevent osteoarthritis or RA and a nutritive or beverage compound or food additive compound comprised of chicken type II collagen prepared by the method of the invention. The invention also involved the use of this nucleic acid molecule in gene therapy.

(57) 摘要

本发明涉及一种包含如 SEQ ID NO: 1 所示编码全长鸡 II 型胶原蛋白的多核苷酸序列之核酸分子, 或其具有相同生物学功能的片段; 及其所编码的鸡 II 型胶原蛋白。本发明还涉及一种制备所述鸡 II 型胶原蛋白的方法, 以及由本发明所述方法制备的鸡 II 型胶原蛋白用于制备治疗和/或预防类风湿性关节炎的药物的用途。本发明特别涉及一种用于预防和/或治疗骨关节炎、类风湿性关节炎的药物组合物, 以及食品或饮料组合物或食品添加剂组合物, 其中含有本发明方法制备的所述鸡 II 型胶原蛋白。本发明还涉及所述核酸分子用于基因治疗的用途。

## 一种编码鸡 II 型胶原蛋白的全长多核苷酸序列及其用途

### 发明领域

本发明涉及一种包含如 SEQ ID NO: 1 所示编码全长鸡 II 型胶原蛋白的多核苷酸序列之核酸分子，或其具有相同生物学功能的片段；及其所编码的鸡 II 型胶原蛋白。本发明还涉及一种制备所述鸡 II 型胶原蛋白的方法，以及由本发明所述方法制备的鸡 II 型胶原蛋白用于制备治疗和 / 或预防类风湿性关节炎的药物的用途。本发明特别涉及一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎(RA)的药物组合物，以及食品或饮料组合物或食品添加剂组合物，其中含有本发明方法制备的所述鸡 II 型胶原蛋白。本发明还涉及所述核酸分子用于基因治疗的用途。

### 发明背景

类风湿性关节炎是严重影响人类健康的常见病、多发病，由于迄今为止对其病因及其发病机制不甚了解，因此控制炎症、缓解症状、维持关节功能仍是目前 RA 治疗的主要手段，远未达到控制关节损伤的目标。近年来，随着分子生物学技术的发展，对 RA 研究取得了巨大进展，尤其是随着对其发病机制认识的不断深入，新的治疗方法和策略也相应产生。

1993 年，美国科学家首次在《科学》(Trentham DE, Dynesius-Trentham RA, Orav EJ, et al. Science, 1993,261: 1727-1730) 报道了他们用天然鸡 II 型胶原蛋白(鸡 CCII)治疗 RA 患者获得巨大成功的研究结果，立刻引起全世界对这一疗法的高度重视。目前美国、英国、法国等发达国家正在进行鸡 II 型胶原蛋白治疗 RA 的 II-III 期临床实验研究。从已发表的研究结

果综合分析可以得出如下基本结论：口服耐受的确为 RA 的治疗开辟了一条能标本兼治极具前景的新途径、新策略和新疗法 (Trentham DE, Dynesius-Trentham RA, Orav EJ, et al., Science, 1993,261: 1727-1730,)。

目前，在鸡 II 型胶原蛋白还未完成 II-III 期临床实验之前，欧美发达国家制药公司纷纷将鸡 II 型胶原蛋白制成食品添加剂，以避开冗长烦琐的药物审批。然而，纵观目前国际上探索口服免疫耐受治疗 RA 所用的均是天然鸡 II 型胶原蛋白，其最大的缺陷就在于：1) 不同公司生产制备的天然鸡 II 型胶原蛋白质量不同，即便是同一公司所生产的不同批号的鸡 II 型胶原蛋白产品质量也会不同。从而也就难以确保疗效的一致性、连续性。2) 天然鸡 II 型胶原蛋白提取和制备所需人力物力甚大。

本发明人提出了解决上述问题的新思路：即采用基因工程方法重组生产鸡 II 型胶原蛋白。为此，提出克隆其编码基因，即 CCOL2A1 基因，并实现高效表达，重组生产鸡 II 型胶原蛋白的设想。本发明就是基于发明人首次克隆成功编码鸡 II 型胶原蛋白的全长多核苷酸序列-CCOL2A1 cDNA 而完成的。

### 发明概述

本发明一个方面涉及一种如 SEQ ID NO: 1 所示包含编码全长鸡 II 型胶原蛋白的多核苷酸序列之核酸分子，或其具有相同生物学功能的片段。

本发明又一方面涉及由本发明所述核酸分子编码的鸡 II 型胶原蛋白，或其具有相同生物学活性的片段。

本发明另外涉及一种含有本发明所述编码鸡 II 型胶原蛋白的核酸分子或其具有相同生物学功能的片段的重组表达载体。

本发明又一方面涉及一种由上述重组表达载体转化的宿主

细胞，其能够表达鸡 II 型胶原蛋白，或其具有相同生物学活性的片段。

本发明再一方面涉及一种制备所述鸡 II 型胶原蛋白的方法，其中包括

- 1).用本发明所述重组表达载体转化合适的宿主细胞；
- 2).在合适的培养基中及适当的培养条件下培养宿主细胞；
- 3).从培养基或细胞中分离、纯化目的蛋白质。

本发明还涉及了本发明所述鸡 II 型胶原蛋白用于制备治疗和 / 或预防类风湿性关节炎的药物的用途。

本发明特别涉及了一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎的药物组合物，其中包含治疗有效量的由本发明所述方法制备的鸡 II 型胶原蛋白，和任选的，药学可接受的载体。

本发明还包括食品或饮料组合物，其特征在于包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。

本发明也包括一种食品添加剂组合物，其中包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。

本发明还包括了本发明所述核酸分子或其片段用于基因治疗的用途。

### 发明详述

口服耐受是自九十年代国际上兴起治疗自身免疫病的特异性免疫疗法之一，也是近年来免疫学研究领域中最活跃、最富有成效的领域之一。它是指口服某种蛋白质抗原，随后再用该抗原进行胃肠外免疫，引起机体对该抗原产生全身性低免疫应答状态。现已知介导口服耐受的主要机制包括自身细胞抑制、克隆无能、克隆清除以及旁观者抑制，这还远不是其机制的全部。其中决定因素是抗原的用量、种类来源、性质、抗原递呈处理的过程以及

宿主的遗传背景和发育程度。

从已发表的研究结果综合分析可以得出如下基本结论：1). 口服耐受的确为 RA 的治疗开辟了一条能标本兼治极具前景的新途径、新策略和新疗法；2). 口服耐受鸡 II 型胶原蛋白的确对相当一部分 RA 患者疗效十分显著；3). 鸡 II 型胶原蛋白的疗效要远好与其它物种来源的 II 型胶原蛋白，因其富含大量的硫酸软骨素 A 和粘蛋白(proteoglycans)，而粘蛋白中硫酸葡糖胺含量最高，它们有着强有力的作用和软骨修复作用；尤其是来源于 6-8 周龄雏鸡胸骨的鸡 II 型胶原蛋白其含量最高；4). 此外，鸡 II 型胶原蛋白中还含有抗氧化作用的粘蛋白，又称之为软骨基质糖蛋白(CMGP)，它能有效地减少氧化作用对软骨细胞的损伤；5). 鸡 II 型胶原蛋白能有效防止蛋白酶对关节软骨的消化破坏，重新编程已被破坏的软骨细胞和细胞因子，从而明显减少炎症的发生；6). 鸡 II 型胶原蛋白能促进软骨细胞及粘蛋白的合成，增加关节滑液及透明质酸的分泌；7). 鸡 II 型胶原蛋白是强有力的抗炎剂和疼痛缓解剂；8). 鸡 II 型胶原蛋白治疗十分安全无任何毒副作用，这是目前任何其它治疗 RA 药物所不能比拟的。

目前，口服 II 型胶原蛋白诱导机体产生免疫耐受已成为有效治疗类风湿关节炎极为重要的新策略。为了提供大量、优质鸡 II 型胶原蛋白用于上述治疗，本发明人摒弃传统的从天然来源中提取、纯化的方法，采用基因工程方法重组生产鸡 II 型胶原蛋白。

为此，本发明人提出克隆 CCOL2A1 基因并高效表达重组生产 CII 的设想，同时还可深化对 CCOL2A1 基因的全面认识与了解。因为迄今为止，国际上虽已成功地克隆了人、犬、鼠 COL2A1 基因及并进行了染色体上的定位，但对于具有如此重大药用价值、治疗价值和经济价值的 CCOL2A1 基因尚未有克隆成功的报道。亦即，有关 CCOL2A1 全长 cDNA 与基因组 DNA 的克隆、CCOL2A1 在染

色体上的定位、CCOL2A1 在各组织中的构成性表达、CCOL2A1 相关生物信息学以及重组 CCOL2A1 的表达研究在国际上还是一个空白。

由于编码鸡 II 型胶原蛋白的 cDNA 基因序列较长，而目前对其功能区的 cDNA 序列所知甚微，因此获得编码鸡 II 型胶原蛋白三螺旋区的 cDNA 无疑是非常重要。由于 II 型胶原具有复杂的二级结构，高重复序列，高 GC 含量（平均 GC 含量大于 70%，个别区域 GC 含量高达 80%）(Nah DH, Upholt WB., J Biol Chem, 1991, 266 34:23446-23452)，因此，直接进行 PCR 非常困难。

为此，我们首先采用 Goldkey 软件对已知的编码鸡 II 型胶原蛋白 3' 原肽区 cDNA 序列进行酶切位点分析，尝试采用限制性内切酶不完全消化 cDNA，再用消化后的 cDNA 进行 C 端原肽区的 PCR 扩增的策略获得成功。鉴于编码鸡 II 型胶原蛋白的 cDNA 基因序列较长，我们将 cDNA 分成五部分分别进行 PCR 扩增，并使各扩增片断彼此之间重叠至少 50bp，便于最后采用 SOE-PCR 策略将各 cDNA 片段进行首尾连接。针对编码鸡 II 型胶原蛋白胶原基因中的高 GC 含量，在 PCR 扩增中应选用适宜扩增高 GC 含量且保证性强的 Tag 酶。

最终本发明人获得了长度为 4837bp 的 CCOL2A1 全长 cDNA，它包括 4260bp 开放阅读框和 520bp 的 3' 非翻译区。将其插入 pGEM-T 载体进行序列测定证实，该 cDNA 序列为 CCOL2A1 所特有的迄今未曾报道过的基因序列。序列测定结果登录于 GenBank (AY046949)。

已有的研究表明，鸡 II 型胶原蛋白基因组 DNA 具有复杂的结构。鉴于国际上迄今为止对于 CCOL2A1 基因组 DNA 尚无任何详细研究资料加以阐述这一事实，我们在成功地克隆了编码 CCOL2A1 全长 cDNA 的基础上，随后又对 CCOL2A1 基因组 DNA

进行了克隆分析。由于 CCOL2A1 基因组 DNA 具有含高量 GC、重複序列多、引物非特异结合部位太多等复杂的结构，最初我们通过 PCR 对鸡外周血细胞（EDTA 抗凝）中提取的 DNA 进行扩增，仅能获得 CCOL2A1 3' 端 5494bp 的片段，而对 3' 端以外的基因组序列，无论如何调整模板、引物及 PCR 扩增条件，均未能获得该目的基因。随后，又对从鸡基因组 DNA 文库中筛选到的阳性克隆进行测序，但也因高 GC 含量及 PolyT 结构限制了测序反应的进行，使测序信号中断。

为此，我们采取了利用已知的外显子序列，设计多个测序引物，才使 CCOL2A1 基因组 DNA 的测序得以完成。最终我们获得长度为 12003bp 的 CCOL2A1 基因组 DNA 序列，其中包含 45 个外显子和 44 个内含子。经查新检索证实，该基因组 DNA 序列为 CCOL2A1 基因组所特有的国际上未曾报道过的。相关基因序列亦已递交 GenBank，Accession No 为 AF452711。

通过对 CCOL2A1 基因组 DNA 序列的分析不难看出，与鸡 I 型胶原和 III 型胶原相比，CCOL2A1 基因的内含子与外显子明显的较小，这与 Upholt 等曾报道的相同 (Ausar SF, Beltramo DM, Castagna LF, et al., Rhematol Int., 2001, 20 :138-144)；II 型胶原 (CII) 在进化上高度保守，不同种属的 II 型胶原蛋白其氨基酸同源性均高于 92%，甚至达 99%。因此，我们认为造成 CCOL2A1 基因紧凑的原因并非内含子数量上的减少，而是由于内含子平均长度较小所致。

为了更好地认识和掌握 CCOL2A1 基因生物信息学的相关内容，深入了解 CCOL2A1 基因的进化特征，本研究利用 Dnastar 软件包中的 MegAlign，将我们所克隆的 CCOL2A1 全长 cDNA 及其相应蛋白质序列与从 GenBank 数据库中检索到的犬 (AF023169, AF242201)、人 (L10347)、斑点鱼 (U23822) 和小鼠 (M65161)

等不同种属的 II 型胶原三螺旋区 cDNA 序列及蛋白质序列进行了同源性比较与分析，通过 Genedoc 程序编辑，绘制出进化树。结果显示，鸡 CII 与犬 CII 在种族进化上同源性最高，cDNA 与氨基酸同源性分别为 79.03 % 和 94.77 %；其次为鸡与人的 CII，同源性分别为 78.96% 和 93.89%；鸡与小鼠 CII 的同源性最低分别为 77.38% 和 92.90%。人与犬 CII 在所比较的五种种属中同源性最高，核苷酸与氨基酸同源性分别为 91.89% 和 98.52%；其次为人与鼠 CII，同源性分别为 88.63% 和 96.15%。

在成功地揭示了有关 CCOL2A1 cDNA、基因组 DNA 及其进化特征后，及时准确地将 CCOL2A1 在染色体上进行定位则显得十分重要。虽然 RH 杂交作图是目前基因在染色体定位最精确的方法，但限于鸡 RH 杂交板目前尚无厂家出售。为此我们采用染色体 FISH 技术，利用地高辛标记的 CCOL2A1 3'UTR 基因片段为探针，与从鸡外周血细胞中制备的中期分裂相染色体进行杂交。经统计分析确证，编码 CCOL2A1 的基因定位于 4 号染色体短臂 2 区，而小鼠、人相应的 COL2A1 则分别定位于 8 号及 12 号染色体(Barnett ML, Combitchi D, Trentham DE., Arthritis & Rheumatism, 1996, 39 4:623-628)。

我们知道，胶原在机体内广泛存在，不同的胶原型可公用其中的一条链，XI 型胶原与 II 型胶原之间则属于此情况。XI 型胶原是由三个不同的多肽亚单位 $\alpha_1$ 、 $\alpha_2$ 、 $\alpha_3$  组成的，其中 $\alpha_3$  是 $\alpha_1$ (II) 基因的表达产物，即 $\alpha_3$  与 II 型胶原 $\alpha_1$  链仅羟化赖氨酸含量不同(Rousseau JC, Farjanel J, Boutillon MM, et al., J Biol Chem, 1996, 271(39): 23743-8)；XIII 型胶原为 II 型胶原的跨膜形式，即 XIII 型胶原与 II 型胶原链的组成完全相同(Snellman A, Keranen MR, Hagg PO, et al. J Biol Chem, 2000, 275(12):8936-44)。

为了深入认识 CCOL2A1 在鸡整个胚胎发育进化中的重要作用，本发明采用 RT-PCR 与双抗体夹心 ELISA 法分别对 CCOL2A1 在鸡胚与成鸡各组织中的构成性表达进行了全面系统的分析。结果表明，CCOL2A1 mRNA 在发育鸡胚的心、肝、玻璃体、角膜、皮肤、胸肌、胸骨、小肠、关节软骨、半月板及颅骨中存在着相应表达，而在脾脏、胸腺与睾丸中未检测到 CCOL2A1 mRNA 的表达；对成鸡 CII 蛋白质水平的检测显示，CII 蛋白除了在胸骨与关节软骨中有表达外，在胰脏与小肠中亦有表达，其中以关节软骨中的表达水平为最高，其次为胸骨。应说明的是，在成年鸡 II 型胶原蛋白表达谱的检测中，由于 ELISA 实验中鸡 II 型胶原蛋白经胃蛋白酶、弹性蛋白酶消化后，II 型胶原形成  $\alpha$  链小短肽，因此含有  $\alpha_1$  (II) 链的胶原均可与鸡 II 型胶原蛋白单克隆抗体结合，因此实验中不能排出同时检测到 XI, XIII 型胶原蛋白的可能。

应强调的是，CCOL2A1 外显子的剪切方式十分复杂，尤其是 II 型胶原外显子 2 在不同的组织采取不同的剪接方式。为此，我们对鸡胚发育时的软骨组织如胸骨、关节软骨和非软骨组织如心脏、肝脏、胸肌、小肠、眼球玻璃体、角膜中 N 端外显子 2 剪切方式进行了分析。在实验中我们发现 17 日龄鸡胚胸骨、生长板、关节软骨中获取的编码鸡 II 型胶原蛋白的 cDNA 不含外显子 2；而心脏、肝脏、眼球玻璃体、角膜、小肠、肌肉、皮肤、胸肌等非软骨组织中以含有外显子 2 的表达存在；其中眼球玻璃体、角膜中提取的总 RNA，经 RT-PCR 发现存在含与不含有外显子 2 两种情形；但眼球玻璃体 RT-PCR 显示，在较小日龄的鸡胚（小于 14 天）仅以含外显子 2 的形式存在（数据未给出）；而当鸡胚发育至 17 天，RT-PCR 结果却出现另外一种情形，以不含外显子 2 为主要表达形式。这是迄今国际上首次对 CCOL2A1 外显子 2 在不同组织中进行剪切方式的研究报道。

对于上述现象只能以下面的解释作为答案：鸡 II 型胶原蛋白在软骨组织中已不含外显子 2 的形式存在，而在非软骨组织中如心脏、肝脏、角膜、小肠、肌肉、皮肤等中主要以含有外显子 2 的形式存在；II 型胶原基因组存在调控序列，其调控 II 型胶原在胚胎发育的不同时期及不同组织中的表达；鸡 II 型胶原蛋白在各脏器广泛存在的现象随着鸡的成长发育而消失，而代之于鸡 II 型胶原蛋白仅在软骨组织中表达；发育的鸡胚胸骨中鸡 II 型胶原蛋白含量丰富（Young MF, Vogeli G, Nunez AM, et al., Nucleic Acids Res, 1984,12 (10): 4207-4228; Marshall GE, Konstas AGP, Lee WR., BrJ Ophthalmol, 1993,77:515-524），但随鸡的生长发育，胸软骨逐渐钙化，鸡 II 型胶原蛋白含量逐渐下降。关于鸡 II 型胶原在机体内的表达，目前认为鸡 II 型胶原在成熟的个体存在于眼球玻璃体与软骨（Seery CM, Davision PF., Invest Ophthalmol Vis Sci,1991,32:1540-1550；Huerre-Jeanpierre C, Mattei MG, Weil D, et al., Am J Hum Genet, 1986, 38(1): 26-37），而我们对天然鸡 II 型胶原的表达检测中却发现眼球玻璃体中不存在 II 型胶原的表达。

为实现本发明的目的，本发明人对编码鸡 II 型胶原蛋白的 cDNA 成功进行了克隆、染色体定位、表达谱分析、基因组 DNA 克隆以及 CCOL2A1 相关生物信息学分析。尤其是编码鸡 II 型胶原蛋白全长 cDNA 的克隆成为本发明完成的坚实基础，意义重大。鸡 II 型胶原蛋白作为口服耐受治疗 RA 的新策略，其 cDNA 的克隆无疑将使基因工程方法生产重组鸡 II 型胶原蛋白成为可能。

具体的，本发明一个方面涉及一种如 SEQ ID NO: 1 所示包含编码全长鸡 II 型胶原蛋白的多核苷酸序列之分离的核酸分子，或其具有相同生物学功能的片段。

本发明又一方面涉及由本发明所述分离的核酸分子编码序列所编码的鸡 II 型胶原蛋白，或其具有相同生物学活性的片段或其保守性变体。

本发明中，蛋白质或多核苷酸“变体”是指一种具有一个或多个氨基酸或核苷酸改变的氨基酸序列或编码它的多核苷酸序列。所述改变可包括氨基酸序列或核苷酸序列中氨基酸或核苷酸的缺失、插入或替换。变体可具有“保守性”改变，其中替换的氨基酸具有与原氨基酸相类似的结构或化学性质，如用亮氨酸替换异亮氨酸。变体也可具有非保守性改变，如用色氨酸替换甘氨酸。

“缺失”是指在氨基酸序列或核苷酸序列中一个或多个氨基酸或核苷酸的缺失。

“插入”或“添加”是指在氨基酸序列或核苷酸序列中的改变导致与天然存在的分子相比，一个或多个氨基酸或核苷酸的增加。

“替换”是指由不同的氨基酸或核苷酸替换一个或多个氨基酸或核苷酸。

“生物活性”是指具有天然分子的结构、调控或生物化学功能的蛋白质。

“相似性”是指氨基酸序列之间排列对比时相应位置氨基酸残基的相同或保守性取代的程度。用于保守性取代的氨基酸例如，带负电荷的氨基酸可包括天冬氨酸和谷氨酸；带正电荷的氨基酸可包括赖氨酸和精氨酸；具有不带电荷的头部基团有相似亲水性的氨基酸可包括亮氨酸、异亮氨酸和缬氨酸；甘氨酸和丙氨酸；天冬酰胺和谷氨酰胺；丝氨酸和苏氨酸；苯丙氨酸和酪氨酸。

“分离的”一词指将物质从它原来的环境（例如，若是自然产生的就指其天然环境）之中移出。比如说，一个自然产生的多

核苷酸或多肽存在于活动物中就是没有被分离出来，但同样的多核苷酸或多肽同一些或全部在自然系统中与之共存的物质分开就是分离的。这样的多核苷酸可能是某一载体的一部分，也可能这样的多核苷酸或多肽是某一组合物的一部分。既然载体或组合物不是它天然环境的成分，它们仍然是分离的。

本发明的多核苷酸可以RNA或DNA的形式存在，其中DNA包括cDNA、基因组DNA和合成的DNA。DNA可以是双链或者单链的，单链也可以是有义链或者反义链。编码多肽的多核苷酸序列可以与SEQ ID No.1所示的多核苷酸序列相同，或者可以是一个由于遗传密码的冗余或简并性而不同的多核苷酸序列，但它编码与SEQ ID No.2相同的成熟多肽。

本发明进一步涉及利用本发明所述的多核苷酸的各种变体，它们编码含有SEQ ID No.2所示的推定的氨基酸序列多肽的片段、类似物和衍生物。这些多核苷酸变体可以是天然存在的等位基因变体或非天然存在的多核苷酸变体。

编码本发明多肽的多核苷酸也可能与特定的标记序列在同一读框中相融合，标记序列可帮助本发明多肽的纯化。标记序列在使用细菌宿主时可以是pQE载体的六聚组氨酸标记，以利于融合有标记的成熟多肽的纯化，或者当使用哺乳类宿主（如猴肾成纤维细胞的COS-7细胞系）时，标记序列可以是一种血细胞凝集素（HA）标记。此外，包含编码本发明多肽的多核苷酸序列还可包含同源或异源的特定信号肽序列，以帮助目的蛋白质分泌到原核细胞或真核细胞膜外。本领域普通技术人员知晓，上述标记序列和信号肽序列也可通过重组方法或化学法添加在用于表达本发明多肽的载体上。

本发明进一步涉及具有SEQ ID No.2所示推定氨基酸序列的多肽及其活性片段、类似物和衍生物。

SEQ ID No.2所示多肽的“具有相同生物学活性的片段”指能基本保留该多肽的生物学功能或活性的多肽。

由本发明所述方法制备的鸡II型胶原蛋白为重组蛋白、多肽或其片段、衍生物和类似物。具体的，SEQ ID No.2所示多肽的片段、衍生物和类似物可以是：(i)一个或多个氨基酸残基被保守性或非保守性氨基酸残基所取代（优选是保守性氨基酸残基）的多肽，取代的氨基酸残基是或不是由遗传密码所编码的氨基酸。例如，可以通过氨基酸残基的插入、取代和/或删除，得到鸡II型胶原蛋白的沉默突变体或功能等同物。可基于氨基酸残基之间在极性、电荷、溶解性、疏水性、亲水性和/或两亲性等方面相似性进行保守性氨基酸取代，只要保留鸡II型胶原蛋白的活性；或(ii)一个或多个氨基酸残基带有取代基团的多肽；或(iii)成熟多肽与其它功能性化合物，如提高多肽半寿期的化合物（例如聚乙二醇）融合在一起的多肽；或(iv)成熟多肽与其它氨基酸序列相融合的多肽，其中所述其它氨基酸序列包括诸如帮助纯化成熟蛋白的氨基酸序列或蛋白原序列等。这些帮助纯化的结构域包括但不限于：金属螯合肽，如用于在固定化金属上纯化的组氨酸-色氨酸模块；用于在固定化免疫球蛋白上纯化的蛋白A结构域以及用于FLAGS延伸/亲和纯化系统的结构域( IMMUNEX公司, Seattle, Wash.)。特异于因子XA或肠激酶的断裂接头序列也可用于帮助目的蛋白质的纯化 (Porath, J. 等人(1992), Prot.Exp.Purif. 3:263-281)。从这些公开内容看，这样的片段、衍生物和类似物的应用于本领域技术人员的知识范围内。

本发明的鸡II型胶原蛋白包括SEQ ID No.2所示的鸡II型胶原蛋白，即成熟多肽，也包括与SEQ ID No.2多肽至少有至少有90%相似性（优选是90%的相同性）的多肽，更优选至少有95%相似性（优选是95%相同性）的多肽，也包括这些多肽的一些部

分，通常这些多肽部分至少含有30个氨基酸、优选至少50个氨基酸。

本发明的多肽、其保守性变体和生物活性片段及衍生物可以用常规的肽合成的方法制备，例如固相肽合成（Merrifield J. (1963), 美国化学学会杂志 (J. Am. Chem. Soc.) 85:2149-2154; Roberge, J. Y. 等人, (1995) 科学, 269:202-204）。蛋白质合成可以手工完成，也可利用肽自动合成仪如 Applied Biosystems 431A 肽合成仪进行 (Perkin Elmer)。本发明鸡II型胶原蛋白也可以从天然生物材料中经分离纯化得到，但优选利用本发明提供的多核苷酸用重组DNA技术制备。

根据常规的重组 DNA 技术，利用本发明的多核苷酸序列可表达或制备重组的鸡 II 型胶原蛋白。本发明因而涉及制备本发明鸡 II 型胶原蛋白的方法，一般包括以下步骤：

- (1).用本发明编码鸡 II 型胶原蛋白的多核苷酸（或变体）或含有该多核苷酸的重组表达载体转化合适的宿主细胞；
- (2).在合适的培养基中及适当的培养条件下培养宿主细胞；
- (3).从培养基或细胞中分离、纯化目的蛋白质。

本发明也涉及包含本发明多核苷酸的重组载体、带有本发明重组载体的遗传工程化宿主细胞和通过重组技术制备本发明多肽的方法。

本发明的多核苷酸可以用来经重组技术产生多肽。例如，该多核苷酸可以存在于选自多种用于表达多肽的表达载体中的任一载体上，这些载体包括染色体的、非染色体的及合成的DNA序列，例如，SV40的衍生物、细菌质粒、噬菌体DNA、酵母质粒、衍生于质粒和噬菌体DNA结合的载体、病毒DNA、杆状病毒、牛痘病毒、腺病毒、禽痘病毒及伪狂犬病毒，包括但不限于pQE系列(Qiagen)、pBS、pD10、phagescript、psiX174、pBluescript SK、

pBSKS、pNH系列 (Stratagene)、pTRC99a、pKK223-3、pDR540、pRIT5 (Pharmacia); pWLNEO、pSV2CAT、pOG44、pXT1、pSG (Stratagene)、pSVK3、pBPV、pMSG、pSVL (Pharmacia)。不过，只要是能在宿主中复制和存活，其它质粒或载体也可应用。

本发明也包括含有本发明多核苷酸的重组构建体。该构建体包括载体，如上述质粒或病毒载体，其中可正向或反向插入本发明的多核苷酸序列。构建体中还包含调节序列，例如，有效地连接到本发明多核苷酸序列上的启动子（包括组成型和诱导型启动子），介导下游结构序列的转录。合适的启动子包括但不限于， $\lambda$ 噬菌体的PL启动子、杆状病毒多角体蛋白启动子；细菌启动子如；LacI、LacZ、T3、T7、gpt、 $\lambda$  PR、PL和trp；真核启动子如：CMV立即早期启动子、HSV胸苷激酶启动子、早期和晚期SV40启动子、逆转录病毒的LTR和小鼠金属硫蛋白I启动子；还包括衍生自植物细胞基因组中的启动子，如热休克蛋白、RUBISCO启动子。

表达载体也包含翻译起始所需要的核糖体结合位点和转录终止子，其还可以具有增强表达的合适序列，如增强子。增强子是DNA的顺式作用因子，通常有大约10-300bp，作用于启动子，提高它的转录。例如，SV40中位于复制起点后侧100-270bp处的增强子，巨细胞病毒早期启动子增强子、位于复制起点后侧的多形瘤增强子及腺病毒增强子。此外，表达载体优选还包含能提供表型特征的一种或多种选择性基因、抗性基因和/或标记基因，以便于转化宿主细胞的筛选。选择性基因如帮助细胞利用吖啶或组氨醇的trpB、hisD (Hartman, S. C., 和 R.C.Mulligan(1988) 美国国家科学院院报 85:8047-51)，用于tk<sup>-</sup>或aprt<sup>-</sup>细胞的肝疱疹病毒胸腺嘧啶激酶 (Wigler, M. 等人(1977) 细胞 11:223-32) 和腺嘌呤磷酸核糖转移酶 (Lowy, I. 等人 (1980) 细胞

22:817-23); 抗性基因如赋予氨甲蝶呤抗性的二氢叶酸还原酶 DHFR ( Wigler, M. 等人(1989), 美国国家科学院院报, 77:3567-70 )或赋予新霉素和G-418抗性的npt( Colbere-Garapin, F.等人(1981) 生物化学杂志 ( J.Mol.Biol. ), 150:1-14 ), 以及四环素或氯苄青霉素抗性基因。哺乳类表达载体一般包括复制起点、适当的启动子、增强子及任何必需的核糖体结合位点、多腺昔化位点、拼接供体和受体位点、转录终止序列和5'侧翼非转录序列。衍生于SV40拼接序列的DNA序列和多腺昔化位点可用于提供所需要的非转录遗传元件。此外, 包含编码本发明多肽的核苷酸序列的载体还可包含同源或异源的特定信号肽序列, 以帮助目的蛋白质分泌到原核细胞或真核细胞膜外。本领域普通技术人员知晓, 上述表达载体及构建体中含有的标记序列和信号肽序列也可通过重组方法或化学法添加在编码本发明多肽的多核苷酸上。

合适载体和启动子的选择为本领域普通技术人员周知。细菌适用的有效表达载体可以这样来构建: 将编码目的蛋白的结构DNA序列随同适当的翻译起始和终止信号被插入到带有一个功能启动子的可操纵的阅读框中。本领域普通技术人员周知用于构建含有本发明核苷酸序列以及合适的转录及翻译调控元件的方法。这些方法包括体外重组DNA技术、合成技术以及体内遗传重组技术 ( Sambrook, J. (1989), 分子克隆实验室手册, Cold Spring Harbor Press; Plainview, N.Y.; Ausubel, F.M. (1989) 当代分子生物学方法 ( Current Protocols in Molecular Biology ), John Wiley & Sons, N.Y. )。

包含上述合适的DNA序列、合适的启动子或控制序列的载体可以用于转化合适的宿主, 让宿主表达该蛋白。

本领域技术人员知晓, 根据本发明DNA序列所插入的表达载体或构建体的种类和特性选择合适的宿主以表达目的蛋白质。适

于表达本发明的多肽的宿主包括但不限于：原核宿主，诸如大肠杆菌、芽孢杆菌属、链霉菌属等；真核宿主，诸如：酵母属、曲霉属、昆虫细胞，諸如果蝇S2和草地夜蛾Sf9；动物细胞，如CHO、COS（猴肾成纤维细胞系，Gluzman (细胞 23:175, 1981) 及其它的能表达相容载体的细胞系，例如C127、3T3、CHO、HeLa、BHK、Bowes黑素瘤细胞；植物细胞以及腺病毒等等。各种哺乳动物细胞的培养系统也能用于表达重组蛋白。从这些讲授看，合适宿主的选择应该在本领域技术人员的知识范围内。

带有如上所述的含有本发明核苷酸序列之载体或构建体能够通过传统的方法导入合适的宿主细胞中以产生重组产物。将构建体导入上述宿主细胞的方法为本领域技术人员周知，包括但不限于：氯化钙介导的转化、磷酸钙转染、DEAE-葡聚糖介导的转染、电穿孔、显微注射、粒子轰击法或基因枪方法 (Sambrook, J.(1989), 分子克隆实验室手册, Cold Spring Harbor Press; Plainview, N.Y.; Ausubel, F.M. (1989) 当代分子生物学方法, John Wiley & Sons, N.Y.; Hobbs, S. 等人, McGraw Hill Yearbook of Science and Technology (1992), McGraw Hill, N.Y.191-196; Engelhard, E.K.等人, 美国国家科学院院报, 91:3224-3227; Logan, J.等人, 美国国家科学院院报, 81:3655-3659)。

在适当的培养条件与培养基中培养经转化的宿主菌株或细胞，使其生长到恰当的细胞密度之后，用适当的方法（例如温度转变或化学药品诱导）诱导所选择的启动子，并将细胞再培养一段时间。针对不同的宿主菌株或细胞选择以及所表达的目的蛋白质的性质相应的培养条件和培养基在本领域技术人员知识范围之内。

在合适的启动子控制下可以在哺乳类细胞、酵母、细菌或其它细胞中表达成熟蛋白。利用由本发明的DNA构建体衍生的

RNA，也可以用无细胞翻译体系产生这种蛋白质（Sambrook, J. (1989), 分子克隆实验室手册, 第18章第4节, Cold Spring Harbor Press; Plainview, N.Y.）。

通常用离心的方法收获细胞或培养液。对于目的蛋白质保留在胞内的情况，一般可用任何便捷的物理、化学方法或酶法，包括冻融循环、超声波、机械破碎，或使用细胞溶解剂或特定的酶破碎细胞，将所得粗提物保留以进一步纯化。对于目的蛋白质分泌到胞外的情况，可直接从培养上清液中利用常规方法回收目的蛋白质。这些方法都是本领域的技术人员熟知的。

将多肽从重组细胞培养物中回收和纯化的方法有硫酸铵或乙醇沉淀、酸提取、阴离子或阳离子交换层析、体积排阻层析、疏水相互作用层析、亲和层析、羟基磷灰石层析和植物凝集素层析。形成成熟蛋白的完整构象还需要蛋白质的重折叠步骤。通常高效液相层析 (HPLC) 或毛细管电泳可应用于最后的纯化步骤。

本发明还涉及了根据本发明所述多核苷酸序列制备的鸡 II 型胶原蛋白在制备治疗和 / 或预防类风湿性关节炎等病的药物的用途。具体的，本发明所述核酸分子的编码区或其部分，如全长序列、部分序列或其实变体均能用于有效表达本发明所述鸡 II 型胶原蛋白或其功能性片段。

本发明特别涉及了一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎的药物组合物，其中包含治疗有效量的由本发明所述方法制备的鸡 II 型胶原蛋白，和任选的，药学可接受的载体。

本发明还包括食品或饮料组合物，其特征在于包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。在发明的一个实施方案中，所述鸡 II 型胶原蛋白用于制备保健食品或食品添加剂等产品以对尤其是对骨关节病中的类风湿性关节炎、骨关节炎等病起到保健作用。

本发明也包括一种食品添加剂组合物，其中包含一定量的由本发明所述方法制备的鸡 II 型胶原蛋白。

本发明还包括了本发明所述核酸分子或其片段用于基因治疗的用途。

#### 附图说明

图 1。显示本发明所述鸡 II 型胶原蛋白编码基因的 PCR 克隆策略。

图 2。显示本发明所述全长鸡 II 型胶原蛋白的结构。

图 3. 显示鸡胚中 II 型胶原蛋白编码基因 3 非翻译区(UTR)的 PCR 结果。

图 4. 显示成鸡中 II 型胶原蛋白编码基因的组织特异性表达。

图 5. II型胶原蛋白浓度与吸光度标准曲线

图 6. 鸡II型胶原蛋白编码基因外显子的剪切分析

图 7. 鸡II型胶原蛋白编码基因的染色体分带。

图 8. 鸡染色体中期分裂相

图 9. ISH杂交结果。

图 10. 人、犬、鼠、鸡、斑点鱼之II型胶原蛋白与基因同源性比较进化树。

图 11. pPIC9K/CCOL2A1 (A) 与及 pPICZ $\alpha$ B/CCOL2A1 (B) 诱导表达上清及胞浆中表达产物的 WB 分析，其中，M 为蛋白分子量标准，1-5 为 WB 阳性带，6 为空载体对照。

图 12 pPICZ $\alpha$ B/CCOL2A1/GS115 与 pPIC9K/P4H $\alpha$ . pPIC9/P4H $\beta$  共表达的 WB 分析，其中，M 为蛋白分子量标准，1-5 为 WB 阳性带，6 为空载体对照。

本发明将在下面的实施例中进一步描述。但是本发明并不局

限于这些实施例。

其中，分子生物学、生物化学、免疫学等常规操作均按照“分子克隆：实验指南”第2版中的描述进行。所述酶切条件和缓冲液均按照生产厂家的说明书操作。通常所用的温育时间大约是37℃一个小时，但根据厂家的指示可以有所改变。

### 实施例

#### 实施例1 鸡胚胸骨总RNA提取

17日龄SPF鸡胚（中国农业科学院畜牧研究所，北京）经消毒后破壳，取鸡胚，置冰生理盐水中冲洗，无菌取胸骨，迅速置Trizol（Invitrogen）中快速研磨，转移至1.5ml eppendorf管，室温放置10min，4℃、10000g离心10min，将上清转移至一DEPC处理的eppendorf管，每管加200μl氯仿，混匀，室温放置10mins，离心同上。将水相转移至一新eppendorf管，加等量异丙醇，室温再放置10min，离心，75%乙醇洗涤沉淀，无RNase水溶解，-80℃保存待反转录。

#### cDNA合成

提取的总RNA，以3'RACE试剂盒（TAKARA）中的oligodT-3sites adaptor primer作为下游引物进行反转录。反转录完毕95℃处理5分钟灭活反转录酶并将cDNA置-20℃保存备用。

#### 鸡COL2A1 cDNA C-端原肽cDNA的PCR扩增

根据已知鸡COL2A1 C端UTR的序列（Sandell LJ, Prentice HL, Kravis D, Upholt WB., J Biol Chem, 1984,259(12) 7826-7834），利用引物col2a1-2F, col2a1-2R（如表1所示），扩增鸡COL2A1 C-端原肽部分。C-端原肽(COL2A1-2) cDNA进行

以下 PCR 扩增程序：引物 col2a1-2F, col2a1-2R, 10 × PCR 缓冲液；5 μL, MgCl<sub>2</sub> (25Mm) 5 μL, dNTP Mix (10Mm) 1 μL, 引物各自为 1 μL (20pmol/μL) 1 μL, 100% 甘油 2.5 μL (终浓度 5%), cDNA 2 μL, Taq 0.8 μL, 总体积 50 μL。PCR 所用程序 96℃ 10min, 96℃ 1min, 62℃ 1 min, 72℃ 3min, 72℃ 7min。PCR 反应体系内含终浓度 5% 甘油, 1% 琼脂糖电泳检查扩增结果, 并采用凝胶回收试剂盒(Invitrogen)回收目的基因, 连接 pGEM-T 载体 (Promega), 转化 DH5 $\alpha$ , 酶切鉴定后测序。

### 鸡 COL2A1 3'UTR cDNA 的克隆

根据已知鸡 COL2A1 基因序列设计扩增鸡 COL2A1 3'UTR 的上游引物 col2a1-1F (参见表 1)，采用 3'RACE 策略，以反转录相对应的 3 sites adaptor primer (col2a1-1R) 扩增 3'UTR, 其中, 采用 PCR 缓冲液, 引物 col2a1-1R 为 3sites adaptor primers, 采用 96℃ 5min, 96℃ 1min, 63℃ 1 min, 72℃ 30 sec (4 个循环), 96℃ 30sec, 61℃ 30sec, 72℃ 20sec (26 个循环), 72℃ 7min。回收及转化均同前述步骤。

### 鸡 COL2A1 N-原肽 cDNA 的克隆

col2a1-5F, col2a1-5R 为上下游引物扩增 5'N 原肽。采用 PCR 缓冲液, 退火与延伸在同一温度点上。96℃ 5min, 96℃ 1min, 72℃ 50 sec (30 个循环), 72℃ 7min。

### 重组质粒载体的构建及序列测定

质粒提取、酶切、回收、连接、转化、构建重组质粒的鉴定等均按《分子克隆实验指南》(Sambook J, Fritsch EF, Maniatis T. Molecular cloning:a laboratory manual. 2<sup>nd</sup> ed Cold Spring

Harbor Laboratory Press, 1989.) 进行，将上述构建好的鸡 COL2A1 cDNA 系列片段基因以及 SOE 连接的鸡 COL2A1 cDNA 系列片段及全长鸡 COL2A1 分别插入 pGEM-T 载体( Promega )，酶切鉴定后测序。

其中所述质粒的构建和制备如下进行：

凝胶快速提取试剂盒回收过程：PCR 扩增产物经 1% 琼脂糖电泳，切胶回收目的基因，再经凝胶快速提取试剂盒回收。过程如下：

- 1) 含目的基因的凝胶置于 eppendorf 管中，加入 1ml 溶液 L1, 50 水浴 15 分钟，使凝胶彻底溶解；
- 2) 将溶解好的凝胶加入离心柱中，12000rpm 离心，柱内再加入 500μl L1, 室温放置 1 分钟，离心同上；
- 3) 加 L2 700μl, 室温放置 5 分钟，离心同上；
- 4) 12000rpm 再离心 2 分钟，使酒精挥发；
- 5) 将柱转移至 1 新 eppendorf 管中，加入热（约 70 ）的 TE 30μl, 室温放置 2 分钟，12000rpm 离心，柱下的液体即为回收的目的 DNA. 回收 DNA 经 1% 琼脂糖电泳，确定浓度为 40μg/μl.

目的 DNA 与 pGEMT 载体的连接：连接体系如下：2×连接缓冲液 5μl , 目的 DNA 3μl, T 载体 1μl, T4DNA 连接酶 1μl, 共计 10μl, 混匀，4 放置过夜。

大肠杆菌 DH5 $\alpha$  感受态的制备参照 Sambook J, Fritsch EF, Maniatis T. Molecular cloning:a laboratory manual. 2<sup>nd</sup> ed Cold Spring Harbor Laboratory Press, 1989. 中所述方法。

PCR 连接产物转化 DH5 $\alpha$  感受态：取如上述连接好的 pGEMT-col2a1-1 ( 2, 3, 4, 5 等 ) 4μl, 加入 100μl 新鲜制备的

感受态，冰浴 30 分钟，42 热休克 90 秒，再迅速置冰水中 2-3 分钟，加入 900 $\mu$ l 无氨苄 LB，37 °C、160rpm 轻摇 45 分钟，加入 1M IPTG 4 $\mu$ l，16 $\mu$ l (50mg/ml) X-gal，涂布氨苄 LB 平板，37 °C 下培养 16 小时至菌落大小合适为止。

阳性转化子的筛选：挑取 IPTGX-gal 氨苄青霉素 LB 平板上的白色克隆接种含有 200 $\mu$ g/ml 氨苄青霉素的 LB 中，37 °C，200rpm 摆床培养 12 小时，提取质粒。

由上述方法制备得到本发明所述 pGEMT-col2a1-2, pGEMT-col2a1-1, pGEMT-col2a1-3, pGEMT-col2a1-(1+2) 和 col2a1-5, col2a1-4 质粒。

质粒采用 Not I 与 Nco I 双酶切鉴定阳性克隆。Not I 与 Nco I 双酶切体系：10×1 $\mu$ l, 0.1%BSA 1 $\mu$ l, 采用 Promega A7500 试剂盒提取质粒 8 $\mu$ l, Not I 1 $\mu$ l, Nco I 1 $\mu$ l, 水 8 $\mu$ l, 共计 20 $\mu$ l, 37 °C 酶切 2 小时。酶切完毕，1% 琼脂糖电泳鉴定阳性克隆，将酶切阳性的克隆送样测序。核酸测序仪为 PE 公司的 ABI377。

### 鸡 COL2A1 全长 cDNA 的 SOE-PCR

为获得鸡 COL2A1 全长 cDNA，本研究采用重叠延伸-PCR (SOE-PCR) 策略来进行构建 (Horton RM, Hunt HD, Ho SN, et al., Gene, 1989, 77:61-68)，这需进行四次 PCR 反应。

第一次 PCR 反应是以质粒 pGEMT-col2a1-2, pGEMT-col2a1-1 PCR 产物为模板，以 col2a-2F、col2a1-1R 为上、下游引物，使 col2a1-2、col2a1-1 连接起来形成产物 col2a1-(2+1)；采用 Pfu DNA 聚合酶，PCR 扩增条件：96 °C 5min, 96 °C 1min, 60 °C 1 min, 72 °C 150sec (4 个循环)，96 °C 30sec, 58 °C 30sec, 72 °C 130sec (26 个循环)，72 °C 7min.；

第二次 PCR 反应是以质粒 pGEMT-col2a1-3, pGEMT-col2a1-(1+2) PCR 产物为模板, 以 col2a-3F、col2a1-1R 为上、下游引物。PCR 扩增条件: 96℃ 5min, 96℃ 1min, 64℃ 1 min, 72℃ 120 sec (4 个循环), 96℃ 30sec, 62℃ 1min, 72℃ 90sec (26 个循环), 72℃ 7min。

第三次 PCR 反应中, 以 col2a-5F、col2a1-4R 为上、下游引物, 将 col2a1-5、col2a1-4 质粒 PCR 产物为模板。PCR 扩增条件: 96℃ 5min, 96℃ 1min, 62℃ 2 min, 72℃ 120 sec (4 个循环), 96℃ 1min, 60℃ 1min, 72℃ 90sec (26 个循环), 72℃ 7min。

第四次 PCR 反应则以 col2a-5F、col2a1-1R 为上、下游引物, 以前三次连接好的 PCR 产物 (col2a1-5+4, col2a1 1+2+3) 为模板, 可获得全长 col2a1 cDNA。PCR 扩增条件: 96℃ 5min, 96℃ 1min, 64℃ 2 min, 72℃ 210 sec (4cycles), 96℃ 1min, 62℃ 1min, 72℃ 180sec (26 cycles), 72℃ 7min。

鸡 COL2A1 全长 cDNA 克隆所用引物序列如表 1 示:

表 1. 鸡 COL2A1 全长 cDNA 扩增引物序列

|            |   |
|------------|---|
| col2a1-1F  | 5'- TCT ATC GCG CAC CCG TTG TGC -3'           |
| col2a1-1R  | 5'- GTC TTG TAG TGC TAC GGC TTG C -3'         |
| col2a1-12F | 5'- TTG CAG ATG TCT CCA ATA CCA G -3'         |
| col2a1-2R  | 5'- GCA CAA CGG CTC GGG CAA TGT GCT AAC G -3' |
| col2a1-3F  | 5'- GCT CGG AAG CAA CGG CCT CG -3'            |
| col2a1-3R  | 5'- CTC GTC CCG GAC GCG ACG G -3'             |
| col2a1-4F  | 5'- CGC TGC GAT CGT CAT GCG G -3'             |
| col2a1-4R  | 5'- GTA GTG ACC CTA CGC CCG AG -3'            |
| col2a1-5F  | 5'- ACG CCG GCT CTC GTG CTC CTC GTG GTG C -3' |
| col2a1-5R  | 5'- CCG CCC GGG TCC GAA TGC CCG CAT -3'       |

本研究采用重叠延伸-PCR (SOE-PCR) 策略来获得全长 4837bp 鸡 COL2A1 cDNA (图 1)。由于鸡 COL2A1 基因 GC 含量较高，基因扩增过程中反应体系采用适宜扩增高 GC 含量的缓冲液，扩增产物特异性极高，将其插入 pGEM-T 载体。利用大肠杆菌 DH5 $\alpha$ ；以及 Promega 出品的工具酶进行酶切、转化和序列测定。其中上述引物由上海博亚生物技术有限公司合成，测序由日本 TAKARA 公司完成。所得结果见后附序列表中 SEQ ID NO: 1。

#### 实施例 2 鸡 COL2A1 基因组 DNA 的提取

鸡 COL2A1 基因组 DNA 的提取采用 Wizard genomic DNA purification kit (Promega) 进行。具体操作见试剂盒说明书。提取 DNA 纯度测定 260/280 为 1.6-1.8 (BECKMAN, DU®640)。

#### 鸡 COL2A1 基因组 DNA 的 PCR 克隆及文库筛选

以 PgF 与 PgR 为上、下游引物，引物序列为：PgF 5' CCA GGC AAG GAT GGC GCA CG 3'；PgR 5'CCT GAT CGG CTC CGC CAA TGT CCA TAG G 3'；进行 CCOL2A1C 端基因组的克隆。选用 LA Taq 酶 (GC buffer) 对 CCOL2A1 基因组 DNA 进行 PCR 扩增。采用 0.8% 琼脂糖电泳检查 PCR 扩增结果，以凝胶回收试剂盒回收目的基因并连至 pGEM-T 载体，酶切鉴定、测序鉴定阳性克隆。由于 CCOL2A1 基因组序列中的高 GC 含量及 polyT 结构，亦通过 SOE-PCR 方法以全血总 DNA 为模板分段克隆 CCOL2A1 基因组 N-端序列，以拼接成接近于全长的 CCOL2A1 基因组 DNA 序列。

最初我们通过 PCR 对鸡外周血中提取的 DNA 进行扩增，能获得 CCOL2A1 3' 端 5494bp 的片段，扩增产物特异性极高 (图 1-9)，

将其插入 pGEM-T 载体，得到 pGEM-T/CCOL2A1，进行测序。

测序结果可以看出鸡 CCOL2A1 3' 端基因组 DNA 包含部分三螺旋区、3' 端肽及 3' 原肽，此克隆基因组部分含有 19 个内含子及 20 个外显子。

然而对 3' 端以外的基因组序列，无论如何调整模板、引物及 PCR 扩增条件，均未能获得该目的基因。随后，又对从鸡基因组 DNA 文库中筛选到的阳性克隆进行测序，但也因高 GC 含量及 PolyT 结构限制了测序反应的进行，使测序信号中断。

为此，我们采取了利用已知的外显子序列，设计多个测序引物，才使 CCOL2A1 基因组 DNA 的测序得以完成。最终我们获得长度为 12003bp 的 CCOL2A1 基因组 DNA 序列，其中包含 45 个外显子和 44 个内含子，内含子、外显子结构如表 2 示。所得鸡 COL2A1 3' 端基因的结构如图 2 所示。经查新检索证实，该基因组 DNA 序列为 CCOL2A1 基因组所特有，国际上未曾报道过。相关基因序列亦已递交 GenBank，Accession No 为 AF452711。

**表 2：鸡 COL2A1 基因组 DNA 3' 端的外显子、内含子结构**

| 外显子 | 编码区氨基酸  | 外显子 sizes (bp) | 内含子 sizes (bp) |
|-----|---------|----------------|----------------|
| 1   | 1-30    | > 135          | ND             |
| 2   | 31-100  | 210            | ND             |
| 3   | 101-117 | 50             | ND             |
| 4   | 118-146 | 87             | ND             |
| 5   | 147-180 | 102            | ND             |
| 6   | 181-206 | 78             | ND             |
| 7   | 207-221 | 45             | ND             |
| 8   | 222-239 | 54             | ND             |
| 9   | 240-258 | 54             | 124            |
| 10  | 259-274 | 54             | 107            |
| 11  | 275-292 | 54             | 100            |
| 12  | 293-310 | 54             | 98             |
| 13  | 311-325 | 45             | 86             |
| 14  | 326-343 | 54             | 110            |
| 15  | 344-358 | 45             | 86             |
| 16  | 359-376 | 54             | 113            |
| 17  | 377-409 | 99             | 89             |
| 18  | 410-424 | 45             | 198            |
| 19  | 425-457 | 99             | 367            |
| 20  | 458-475 | 54             | 885            |
| 21  | 476-511 | 108            | 408            |
| 22  | 512-529 | 54             | 88             |
| 23  | 530-562 | 99             | 91             |
| 24  | 563-580 | 54             | 682            |
| 25  | 581-613 | 99             | 82             |
| 26  | 614-631 | 54             | 82             |
| 27  | 632-649 | 54             | 124            |
| 28  | 650-667 | 54             | 277            |

|             |            |     |     |
|-------------|------------|-----|-----|
| 29          | 668-685    | 54  | 79  |
| 30          | 686-700    | 45  | 114 |
| 31          | 701-733    | 99  | 262 |
| 32          | 734-769    | 108 | 92  |
| 33          | 770-787    | 54  | 80  |
| 34          | 789-805    | 54  | 81  |
| 35          | 806-823    | 54  | 82  |
| 36          | 824-841    | 54  | 94  |
| 37          | 842-877    | 108 | 89  |
| 38          | 878-895    | 54  | 108 |
| 39          | 896-913    | 54  | 113 |
| 40          | 914-967    | 162 | 268 |
| 41          | 968-1003   | 108 | 92  |
| 42          | 1004-1039  | 108 | 309 |
| 43          | 1040-1057  | 54  | 160 |
| 44          | 1058-1093  | 108 | 97  |
| 45          | 1094-1111  | 54  | 268 |
| 46          | 1112-1147  | 108 | 109 |
| 47          | 1148-1165  | 54  | 134 |
| 48          | 1156-1201  | 108 | 79  |
| 49          | 1202-1297  | 289 | 78  |
| 50          | 1298-1360  | 188 | 407 |
| 51          | 1361-1441  | 243 | 594 |
| 52          | 1442-1489+ | 667 | 112 |
| 520bp 3'UTR |            |     |     |

注：外显子、内含子的命名采用 Upholt 等 J Biol Chem, 1984, 259 ( 12 ): 7826-7834 的方法。

### 实施例 3 鸡 COL2A1 $\alpha_1$ ( II ) 组织特异性表达

### 鸡胚 COL2A1 表达谱

COL2A1 3' UTR 经 Blast 的核苷酸序列数据库检索为 COL2A1 所特有。因此以 COL2A1 cDNA 3' UTR 为扩增对象，研究  $\alpha$  1 (II) 在发育的鸡胚中的表达。应用 RT-PCR 对 17 日龄鸡胚心、肝、玻璃体、角膜、皮肤、胸肌、胸骨等脏器进分析，所用引物为表 1 中的 col2a1 - 1F, col2a1 - 1R, 产物连 pGEMT-easy 载体，酶切后测序。同时以 GAPDH 作为内参，利用引物 PF<sub>GAPDH</sub> 5' GC AGA GGT GCT GCC CAG AAC 3'; PR<sub>GAPDH</sub> 5' TCA CTC CTT GGA TGC CAT GTG 3' 扩增 412bp 片段的 GAPDH。

从 17 日龄鸡胚 COL2A1 3' UTR RT - PCR 结果可以看出，鸡 COL2A1 mRNA 在鸡胚的心、肝、玻璃体、角膜、皮肤、胸肌、胸骨、小肠、关节软骨、半月板、颅骨中有表达，而脾脏、胸腺、睾丸却检测不到 COL2A1 mRNA。(参见图 3)。

### 成鸡 II 型胶原蛋白表达谱

采用天然 II 型胶原检测试剂盒 (Chondrex) 对四周龄鸡心、肝、脾、肾、玻璃体、角膜、皮肤、胸肌、半月板、胰脏、胸腺、小肠、胃、睾丸、骨骼肌、大脑、小脑、关节软骨、胸骨、肺脏共 20 个脏器进行 ELISA 检测 II 型胶原蛋白的存在。(冻干组织各 5mg, 充分匀浆, 0.8ml 50mM 醋酸-0.2M NaCl (pH2.9-3.0) 溶解, 100 $\mu$ l 20mg/ml 的胃蛋白酶 (Sigma) 4℃消化 48h; 加入 250 $\mu$ l 10 × TSB, 并用 1 M NaOH 调 pH 至 8.0, 加入 100 $\mu$ l 2mg/ml 的弹性蛋白酶 (Sigma) 混匀, 4℃消化过夜。10000g, 4℃离心 5 分钟, 弃沉淀, 上清定容至 1ml, 4℃保存备用。

采用双抗体夹心 ELISA 法检测提取的 II 型胶原蛋白。选用

两种针对鸡 II 型胶原蛋白分子不同表位的单克隆抗体，其中，一种作为包被抗体，另一种作为酶标检测抗体。首先捕获抗体包被 ELISA 板，4℃过夜，洗涤三次，处理好的样品 1: 200 稀释并与标准品一同点样于包被好的 ELISA 板，37℃ 2h 洗涤同上；生物素标记的鸡 II 型胶原蛋白单克隆抗体加入各孔，37℃孵育 2h，洗涤同上，最后加入酶标链亲和素，37℃1h；充分洗涤，最后加显色液 OPD，37℃显色半小时，1.25M 硫酸终止反应。测 OD490 光密度值。具体操作见试剂盒说明书。

成鸡 20 种脏器天然 II 型胶原蛋白检测实验发现，鸡 II 型胶原蛋白在胸骨、关节软骨、胰脏、小肠有表达，而在其它脏器却不见表达。（参见图 4）

同时以鸡天然 II 型胶原蛋白作为标准品，制作浓度、吸光度标准曲线，依据标准曲线计算相应各阳性组织中鸡 II 型胶原蛋白的表达量（参见图 5）。

据标准曲线得出胸骨中鸡 II 型胶原蛋白的表达量为 0.4%，关节软骨表达量在四周鸡所检测组织中表达量最高达 1.5%，胰脏、小肠中的表达量分别为 0.84% 与 0.2%。

#### 实施例 4 鸡 COL2A1 5'原肽外显子 2 剪接分析

采用 RT-PCR 分析鸡 COL2A1 外显子 2 在鸡胚各脏器中的剪接情况，以表 1 中 col2a1—5F, col2a1—5R 为上、下游引物对鸡皮肤、肝脏、眼球玻璃体、角膜、胸肌、小肠、胸骨、关节软骨等进行 PCR 扩增分析，1% 琼脂糖电泳检测 PCR 扩增产物的特异性。RT-PCR 结果表明，鸡 COL2A1 5'N 原肽外显子 2 存在于鸡心脏、肝脏、眼球玻璃体、角膜、胸肌、小肠，而软骨组织如胸骨、关节软骨中以不含外显子 2 的表达形式存在。（参见图 6）

### 实施例 5 鸡 COL2A1 基因染色体的定位

为了将鸡 COL2A1 基因精细定位于染色体,采用染色体 ISH 技术,应用地高辛高效标记检测试剂盒( Roche )标记的鸡 COL2A1 3'UTR 基因片段为探针,与中期分裂相的鸡染色体进行杂交,研究其在染色体上的定位。

采用 ISH 原位杂交,以地高辛标记的鸡 COL2A1 3'UTR 基因片段为探针,与鸡中期分裂相的染色体精细杂交,杂交信号经统计分析,将鸡 COL2A1 基因定位于 4 号染色体短臂 2 区(参见图 7, 8, 和 9)。

### 实施例 6 鸡 COL2A1 同源性比较

利用 Dnastar 软件包中的 MegAlign 对已获得的鸡 COL2A1 与从 GenBank 数据库中检索到的犬 (AF023169, AF242201)、人 (MM001844)、斑点鱼 (U23822)、小鼠 (M65161) 等不同种属的 II 型胶原三螺旋区 cDNA 序列及蛋白质序列进行同源性的比较和分析,Genedoc 程序编辑,绘制进化树(参见图 10)。

鸡 COL2A1 cDNA 序列与已知的人、犬、小鼠、斑点鱼进行了同源性比较,结果显示鸡 COL2A1 与犬 CII 在种族进化上同源性最高,其氨基酸与 cDNA 同源性分别为 79.03 %, 94.77 %;其次为鸡与人,同源性分别为 78.96%, 93.89%;鸡与鼠同源性分别为 77.38%, 92.90%。人与犬在所比较的五种种属中同源性最高,氨基酸与蛋白质分别为 91.89%、98.52%。其次为人与鼠,分别为 88.63%, 96.15%, 比较结果建立进化树。

### 参考文献

1. Jenkins JK, Hardy KJ. Biological modifier therapy for the treatment of

rheumatoid arthritis. Am J Med Sci, 2002,323(4):197-205.

2. Danos O, Malligan MC. Safe and efficient generation of recombinant retrovirus with amphotropic host ranges. Proc Natl Acad Sci USA,1988,85:6460-6465.
3. Roessler BJ, Allen ED, Wilson JM. Adenoviral-mediated gene transfer to rabbit synovium in vivo. J Clin Invest,1993,92:1085-1092.
4. Trentham DE, Dynesius-Trentham RA, Orav EJ, et al. Effects of oral administration of type II collagen on rheumatoid arthritis.Science, 1993,261: 1727-1730,
5. Sandell LJ, Prentice HL, Kravis D, Upholt WB. Structure and sequence of the chicken type II procollagen gene . J Biol Chem , 1984,259 (12) 7826-7834.
6. Horton RM, Hunt HD, Ho SN, et al. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene,1989, 77:61-68.
7. Sambook J, Fritsch EF, Maniatis T. Molecular cloning:a laboratory manual. 2<sup>nd</sup> ed Cold Spring Harbor Laboratory Press,1989.
8. Nah DH, Upholt WB. Type II collagen mRNA containing an alternatively spliced exon predominates in the chick limb prior to chondrogenesis. J Biol Chem,1991, 266 34:23446-23452.
9. Rousseau JC, Farjanel J, Boutillon MM, et al. Processing of type XI collagen. Determination of the matrix forms of the alpha 1 (XI) chain. J Biol Chem, 1996,271(39): 23743-8.
10. Snellman A,Keranen MR, Hagg PO, et al. Type XIII collagen forms homotrimers with three triple helical collagenous domains and its association into disulfide-bonded trimers is enhanced by proly 4-hydroxylase. J Biol Chem, 2000, 275(12):8936-44.
11. Young MF, Vogeli G, Nunez AM, et al. Isolation of cDNA and genomic DNA clones encoding type II collagen. Nucleic Acids Res, 1984,12 (10): 4207-4228.
12. Marshall GE, Konstas AGP, Lee WR. Collagens in ocular tissues. BrJ

**Ophthalmol, 1993,77:515-524.**

**13. Seery CM, Davision PE.Collagen of the bovine vitreous. Invest Ophthalmol Vis Sci,1991,32:1540-1550.**

**14. Huerre-Jeanpierre C, Mattei MG, Weil D, et al. Further evidence for the dispersion of the human fibrillar collagen genes. Am J Hum Genet, 1986, 38(1): 26-37.**

**15. Ausar SF, Beltramo DM, Castagna LF, et al. Treatment of rheumatoid arthritis by oral tolerance of bovine tracheal type II collagen. Rhematol Int , 2001, 20 :138-144.**

**16. Barnett ML, Combitchi D, Trentham DE. A pilot trial of oral type II collagen in the treatment of juvenile rheumatoid arthritis. Arthritis & Rheumatism, 1996, 39 4:623-628.**

**17. Kim WU, Lee WK, Ryoo JW, et al. Suppression of collagen-induced arthritis by single administration of poly(latic-co-glycolic acid) nanoparticles entrapping type II collagen: a noval treatment strategy for induction of oral tolerance. Arthritis Rheum, 2002, 46:1109-20.**

## 实施例 7 重组 CCOL2A1 全长 cDNA 的表达

实施例中所采用的材料、试剂和仪器如下所述：

pPIC9K、pPIC9、pPICZ $\alpha$ A、B、C 表达载体系统及 GS115、X33 酵母菌株均购自 Invitrogen 公司；绿豆核酸酶购自 NEB 公司；Spheroplast kit for yeast 与 Pichia EasyComp<sup>TM</sup> kit 购自 Invitrogen 公司；Mini II 型蛋白电泳仪及蛋白半干电转仪（Bio-Rad 公司）；GeL-Pro3.1 凝胶成像系统（Media Cybemetic 公司）；DU640 型核酸-蛋白分析仪（Beckman 公司）。

### 1) pPICZ $\alpha$ B/CCOL2A1 真核表达载体的构建

用于构建 pPICZ $\alpha$ B/CCOL2A1 真核表达载体。引物序列 25：5'GGT ACC TTG GTG GAA ACT TTG CGG 3'；引物 26 序列：5'GGT ACC GTT ACA AGA AGC AGA CTG 3'；以质粒 pGEM-T/CCOL2A1 为模板进行 PCR 扩增，PCR 反应体系为：2 × GC Buffer I 25μl, dNTP(2.5mM) 4μl, 引物 25(20μM) 0.5μl, 引物 26(20μM) 0.5μl, PCR 扩增程序如下 96℃ 5 min 4 × (96℃ 1 min, 67℃ 1 min, 72℃ 3 min) 26 × (96℃ 1 min, 65℃ 30 sec, 72℃ 2.5 min) 72℃ 5 min.

回收 PCR 产物克隆入 pPICZ $\alpha$ B 表达载体中，采用限制性内切酶 *Bam*HI 鉴定正确插入方向的表达载体 pPICZ $\alpha$ B/CCOL2A1。

### 2) 酵母表达载体 pPIC9K/CCOL2A1 的构建

为提高 pPICZ $\alpha$ B/CCOL2A1 载体在毕赤酵母中的表达稳定性和表达量，进一步将 CCOL2A1 构建在多拷贝载体 pPIC9K 中，构建 pPIC9K/CCOL2A1。pPIC9K/CCOL2A1 的构建流程详见下图，简述如下：首先 pPICZ $\alpha$ B/CCOL2A1 经 *Kpn* I 线形化后再用绿豆核酸酶进行削平，同时 pPIC9K 经平端酶 *Sna*BI 削平，两者

由 T4 DNA 连接酶连接，构建成 pPIC9K /CCOL2A1 表达载体，经测序确定阅读框正确否。

### 3) *P.pastoris* GS115 原生质体的制备

采用原生质体法转化酵母，流程如下：

- a) 将 *P.pastoris* GS115 划线接种于 YPD 平板，30℃ 培养 2d，挑取单克隆接种于 5ml YPD 培养基中，30℃ 300rpm 培养过夜；
- b) 1: 1000 稀释接种菌悬液至 100ml YPD，30℃ 300rpm 培养至 OD600 为 0.2-0.3 时，1,500g 离心 5min，弃上清；
- c) 分别以 20ml 无菌水、20 ml SED 及 20ml 1M 山梨醇各洗涤细胞一次，每次洗完后 1,500g 离心 5min，弃上清，最后加 20ml SCE 重悬细胞，分成 10ml 的 A、B 两管；
- d) 取 1.5ml eppendorf，每管加 800μl 15%SDS，并分别标上 0、2、4、5、6、8、9、10、15、18、20、25、30、35、40 分别代表 Zymolyase 消化细胞壁的时间 (min)；从 A 管中取 200μl 菌悬液加入后，放入冷冻中作为 Zymolyase 消化分钟的对照管。加 7.5μl Zymolyase (2.25 units) 到 A 管，轻轻混匀，30℃ 保温；根据管上标记的时间，每次从 A 管取 200μl 消化的细胞悬液并加 800μl 5%SDS 至 eppendorf 管中，并立即放置冰上以终止酶反应；
- e) 依据下列公式计算每个时间点形成原生质体的百分比：  
原生质体 % = 100 - [(OD800 (min) / OD800 (0min) × 100)，以计算形成 70% 原生质体的时间(t min)；加 7.5μl Zymolyase 到 B 管，30℃ 保温 t min；
- f) 室温 750g 离心 10min，弃上清，分别用 10ml 1M 的山梨醇 10ml CaS 洗涤原生质体，弃上清，最后用 0.6ml CaS 重悬原生质体，用于转化。

4) pPIC9K/CCOL2A1、pPICZ $\alpha$ B/CCOL2A1 转化 GS115 原生质体

a)、约 10 $\mu$ g 线形化的 pPICZ $\alpha$ B/CCOL2A1、pPIC9K/CCOL2A1 分别与 100 $\mu$ l 刚制备好的原生质体轻轻混合，室温放置 10min；同时以 pPIC9K/GS115 及 pPICZ $\alpha$ B/GS115 空质粒载体作为阴性对照；

b)、加 1ml 新鲜配制的 PEG/CaT (1: 1) 溶液，轻轻混合，室温再放置 10min；

c)、750g 离心 10min，尽可能吸去上清，用 150 $\mu$ l SOS 重悬细胞并静置 20min，加 850 $\mu$ l 1M 的山梨醇；

d)、每 150 $\mu$ l 已转化的原生质体与 10ml RD 琼脂糖培养基混合，并铺在 RDB 平板上层，28-30℃ 培养 4-6d 并观察结果。

5) pPICZ $\alpha$ B/CCOL2A1/GS115 及 pPIC9K/CCOL2A1/GS115 阳性转化子的抗性筛选

培养后的 pPIC9K/CCOL2A1/GS115 与 pPICZ $\alpha$ B/CCOL2A1/GS115 在 RDB 平板长出约 100 个的转化子，为了获得多拷贝，将各转化子分别转移至在含不同浓度的 G418 或 Zeocin 的抗性平板 (G418 浓度：0.25mg/ml、0.5 mg/ml、1 mg/ml、2.0mg/ml、3.0 mg/ml、4.0mg/ml；Zeocin 浓度 100 $\mu$ g/ml, 200 $\mu$ g/ml) 上筛选多拷贝。

6) 酵母基因组 DNA 的提取及目的基因在酵母基因组中整合的 PCR 鉴定

分别将抗性转化子及空载体对照，接种到 2ml 的 YPD 中，30℃ 300rpm 培养 2d，离心后提取基因组 DNA，采用 PCR 检测 CCOL2A1 基因在酵母基因组中的整合，其中以 CCOL2A1 cDNA 的特异性引物：F:5'GGTACC TTG GTG GAA ACT TTG CGG

3'；R: 5'GGTACC GTT ACA AGA AGC AGA CTG 3'。PCR 循环参数为：96℃ 5min → 5 ×(96℃ 1min, 60℃ 2min, 72℃ 3min) → 25 ×(96℃ 1min, 58℃ 1min, 72℃ 2min) → 72℃ 5min。

#### 7) pPICZ $\alpha$ B/CCOL2A1/GS115、pPIC9K/CCOL2A1/GS115 的甲醇诱导表达及免疫印迹分析

将经过鉴定为阳性的克隆，接种至 30ml BMGY 的培养基中 28℃ 300rpm 培养 2d 后离心弃上清，以 1/10 体积的 BMMY 重悬，于 28℃ 300rpm 继续培养；期间每隔 24hr 补充甲醇至 1% 体积。分别收集诱导 4d 培养物的上清及菌体；上清经透析、脱盐后冻干；菌体中加入破菌 buffer 与约 1/4 体积的玻璃珠破菌，10,000rpm 离心，上清即为胞浆液；诱导上清及胞浆液经 10% SDS-PAGE 电泳并转膜分析，20% 胎牛血清封闭，4℃ 过夜；以 1: 1000 稀释的 95D1A 单克隆抗体与 NC 膜作用，室温孵育 2h，TTBS 洗膜三次，每次 5min；然后再用 1: 500 稀释的羊抗小鼠二抗-AP 与 NC 膜室温孵育 2h，洗膜同上；最后用 NBT/BCIP 显色试剂盒显色。

结果：

另取上述各表达载体诱导表达后的上清进行脱盐、浓缩及 10% SDS-PAGE 的电泳分析与 WB 鉴定。结果表明，pPICZ $\alpha$ B/CCOL2A1 转化子中仅在胞浆内可见位于 80KD 左右 CCOL2A1 $\alpha$ 链的表达带（图 11A）；而在 pPIC9K/CCOL2A1 转化子胞浆中 WB 检测到表达完整的 CCOL2A1 $\alpha$ 链（图 11B），上清中却无 WB 条带出现。

#### 实施例 8 脯氨酸 4 羟化酶 $\alpha$ 亚单位和 / 或 $\beta$ 亚单位表达载体的构建

##### 1) 脯氨酸 4 羟化酶 $\alpha$ 亚单位 (P4H $\alpha$ ) cDNA 的克隆

根据鸡脯氨酸 4 羟化酶 $\alpha$ 亚单位 (P4H $\alpha$ ) 基因的序列，设计

引物 a: 5' AGA TAC TGC TAC GAA AGA CCC CGA G 3' ;  
引物 b: 5' CTC TCT TGG TTG TAG CCC TCA TCT G 3' ;  
PCR 反应体系为: 10 ×PCR Buffer 5μl, MgCl<sub>2</sub> (25mM) 5μl,  
dNTP Mix (2.5mM) 4μl, 引物 a (20μM) 0.5μl, 引物 b (20μM)  
0.5μl, cDNA 1μl, Taq 0.5μl, H<sub>2</sub>O 补至 50μl; PCR 扩增产物用  
1% 琼脂糖电泳检测; PCR 扩增程序为: 96℃ 5min → 4 ×(96℃  
1min, 72℃ 3min) → 26 ×(96℃ 1min, 69℃ 1min, 72℃ 2min) → 72  
℃ 3min.

### 2) pPIC9K/P4Hα真核表达载体的构建

将如上述所得 P4Hα的 PCR 产物与 pGEM-T 载体连接, 构建 pGEM-T/P4Hα质粒。通过 PCR 引入 Not I 酶切位点, 所用引物 a: 5' GCG GCC GCA GAT ACT GCT ACG AAA G 3' ; 引物 b: 5' GCG GCC GCC TCT CTT GGT TGT AGG 3' ; 采用 Pfu DNA 聚合酶进行 PCR 扩增, 产物回收后加 A 尾, 连至 pGEM-T 载体内并测序; pGEM-T/P4Hα经 NotI 酶切、回收目的片段, 通过 T4 DNA 连接酶将 P4Hα基因与经 NotI 线形化的酵母 pPIC9K 载体相连接, 构建成 pPIC9K/P4Hα表达载体。鉴定插入方向正确。

### 3) P4HβcDNA 的克隆及重组表达载体的构建

根据已知鸡 P4Hβ亚基基因序列设计引物 F: 5' GCG GCC GCA CAG CCC CTG GAG GAG 3' ; 引物 R: 5' GCG GCC GCG GTG ATG TAG ATC AGT C 3' , 并引入 NotI 酶切位点; 对 17 日龄鸡胚胸骨细胞中提取的 RNA 进行 PCR 扩增, 反应体系中采用 1.5mM 浓度的 [Mg<sup>2+</sup>] , PCR 扩增程序为: 96℃ 5min → 4 ×(96℃ 1min, 53℃ 1min, 72℃ 2min) → 26 ×(96℃ 30sec, 51℃ 30sec, 72℃ 2min) → 72℃ 5min. PCR 产物连入 pGEM-T 载体内进行测序;

测序正确的克隆再导入毕赤酵母 pPIC9 表达载体中，构建成 pPIC9/P4H $\beta$ 表达载体。

#### 实施例 9 pPIC9K/P4H $\alpha$ 、pPIC9/P4H $\beta$ 与 pPICZ $\alpha$ B/CCOL2A1 在酵母 GS115 中的共表达

将如实施例 7 和 8 中所述重组表达载体 pPIC9K/P4H $\alpha$ 、pPIC9/P4H $\beta$ 及 pPICZ $\alpha$ B/CCOL2A1 分别采用 *Bgl*II、*Sal*I、*Pme*I 线形化。各取 10 $\mu$ g 线性化的质粒转化 GS115，在含 G418 及 Zeocin 的 MM 平板筛选阳性转化子，30℃ 培养 4-6d。

#### 实施例 10 pPICZ $\alpha$ A/ P4H $\alpha$ - $\beta$ 双表达载体的构建

pPIC9K/P4H $\alpha$ 、pPIC9/P4H $\beta$ 及 pPICZ $\alpha$ B/CCOL2A1 在共表达时，阳性克隆不稳定，为了克服上述构建的不足之处，将 P4H $\alpha$ 、 $\beta$ 构建在同一表达载体 pPICZ $\alpha$ A，形成 pPICZ $\alpha$ A/P4H $\alpha$ -P4H $\beta$ 双表达载体。

分别以实施例 8 和 9 中所得质粒 pGEM-T/P4H $\alpha$ 及 pGEM-T/P4H $\beta$ 为模板进行 PCR，P4H $\alpha$ 引物 F: 5'GC GGCC CGC GAT ACT GCT ACG AAA G 3'；引物 R: 5'GC GGCC CGC CTC CAA CTC TGA TAA C 3'；P4H $\beta$ 引物 F: 5'GC GGCC CGC CAG CCC CTG GAG GAG -3'；R 5' GC GGCC CGC TTA ATC ATC ATC AGC 3'；PCR 循环参数: 96℃ 5min → 4 × (96℃ 1min, 66℃ 80sec, 72℃ 90sec) → 26 × (96℃ 1min, 64℃ 40sec, 72℃ 50sec) → 72℃ 5min。PCR 扩增产物用 1% 琼脂糖电泳检查，回收特异性扩增目的条带并连接 pGEM-Teasy 载体并测序。

测序正确的 pGEM-T/P4H $\alpha$ 及 pGEM-T/P4H $\beta$ 分别采用 *Not*I 酶切，连接 pPICZ $\alpha$ A，得到 pPICZ $\alpha$ A/P4H $\alpha$ 重组表达载体，和 pPICZ $\alpha$ A/P4H $\beta$ 表达载体。

将 pPICZ $\alpha$ A/P4H $\beta$  完整表达单元切下，插入 pPICZ $\alpha$ A/P4H $\alpha$ 。再根据 *Bgl*II、*Bam*HI 为同裂酶，可连接 *Bgl*II 线性化的 pPICZ $\alpha$ A/P4H $\alpha$ ，构建成 pPICZ $\alpha$ A/P4H $\alpha$ - $\beta$  双表达载体。

根据 P4H $\beta$  表达单元位于 P4H $\beta$  表达单元的上游，应用 PCR 鉴定 pPICZ $\alpha$ A/P4H $\alpha$ - $\beta$  双表达载体中 P4H $\alpha$ 、 $\beta$  表达单元的方向是否一致。上游引物 5' GCGGCCGC CAG CCC CTG GAG GAG 3'；与扩增 P4H $\alpha$  特异性的下游引物 5' GCGGCCGC CTC CAA CTC TGA TAA C 3' 进行 PCR 扩增；如插入为正向，则 PCR 扩增可获得 4644bp 的产物，反之，则 PCR 扩增无目的条带出现。

#### 实施例 11 pPIC9K/CCOL2A1 与 pPICZ $\alpha$ A/P4H $\alpha$ - $\beta$ 共表达

将 pPIC9K/CCOL2A1 与 pPICZ $\alpha$ A/P4H $\alpha$ - $\beta$  共转化 GS115 进行三基因的共表达研究。原生质体的制作、阳性转化子的筛选、PCR 鉴定以及甲醇诱导均按前述方法进行。

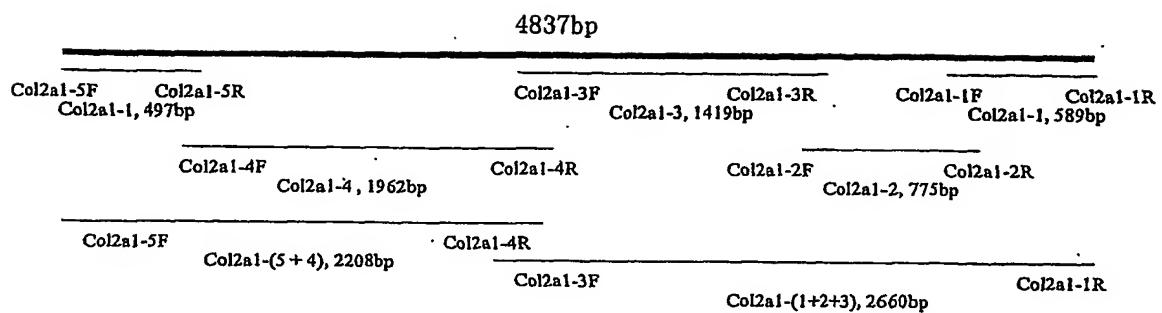
采用 SDS-PAGE 及 WB 法对 pPICZ $\alpha$ B/CCOL2A1/GS115 与 pPIC9K/P4H $\alpha$ 、pPIC9/P4H $\beta$  共表达的诱导上清及胞浆表达物进行鉴定，WB 结果如图 12 示，从图中可看到，共表达时可在其胞浆内得到全长表达的的 CCOL2A1 $\alpha$  链，而在上清中则未检测到表达产物。

## 权 利 要 求

1. 一种如 SEQ ID NO: 1 所示的包含编码全长鸡 II 型胶原蛋白的多核苷酸序列之分离的核酸分子，或其具有相同生物学功能的片段。
2. 权利要求 1 所述的分离的核酸分子，其为具有如 SEQ ID NO: 2 所示编码全长鸡 II 型胶原蛋白的多核苷酸序列片段。
3. 权利要求 1 所述分离的核酸分子或其具有相同生物学功能的片段编码的鸡 II 型胶原蛋白，或其具有相同生物学活性的片段。
4. 一种含有权利要求 1 或 2 所述核酸分子或其具有相同生物学功能的片段的重组表达载体。
5. 一种由权利要求 4 所述重组表达载体转化的宿主细胞，其能够表达鸡 II 型胶原蛋白，或其具有相同生物学活性的片段。
6. 一种制备权利要求 3 所述鸡 II 型胶原蛋白的方法，其中包括
  - 1).用权利要求 4 所述重组表达载体转化合适的宿主细胞；
  - 2).在合适的培养基中及适当的培养条件下培养宿主细胞；
  - 3).从培养基或细胞中分离、纯化目的蛋白质。
7. 权利要求 6 所述方法制备的鸡 II 型胶原蛋白用于制备治疗和 / 或预防类风湿性关节炎药物的用途。
8. 一种用于预防和 / 或治疗骨关节炎、类风湿性关节炎的药物组合物，其中包含治疗有效量的根据权利要求 6 所述方法制备的鸡 II 型胶原蛋白，和任选的，药学可接受的载体。
9. 一种食品或饮料组合物，其特征在于包含一定量的权利要求 6 所述方法制备的鸡 II 型胶原蛋白。
10. 一种食品添加剂组合物，其中包含一定量的权利要求 6 所述方法制备的鸡 II 型胶原蛋白。

11. 一种治疗和 / 或预防类风湿性关节炎的方法，包括向需要的患者施用权利要求 1 所述核酸分子或其具有相同生物学功能的片段。

FIG. 1



鸡 II 型胶原 PCR 克隆策略

## FIG. 2

信号肽

MHGRRPPRSAALLLLLLLTTAAAAA

N-原肽

QDRDLRQPQPKQKGEPGDIKVVGPRGPQGPQGPAGEQGQRGDRGEKGEKGAPGPRGRDGEPGTPGNPGPPGPPGLGGN  
 FAA QMAGGFDEKAGGAQMGMVQ GPMCPMCPRGPPCPTGAPGPQC

N-原肽

螺旋结构域

FQGNPGEPEGAAGPMGPR GPPGPPGKPGDDGETCKPKSGERGPPGPQGARCPGTPGLPGVKGHRGYPGLDGAKG  
 EAGAPGAKGESCGSPGENSPGPMGPRGLPGERGRPGPSGAAGARGNDGLPGPAGPPGPVGPAGAPGFPGAPGSKGEAGPTGARGPE  
 GAQGPRGESGTPGSPGPAGAPGNPGTDGIPGAKGSAGAPGIAGAPGFPGPRCPGPPQGATGPLGPKGQTCEPGIAGFKGEQGPKGE  
 TGPAGPQGAPGPAGEEGKRGARGEPEGAAGPVCPGCGAPGNRGFPQGDLAGPKGAPGERGPAGLAGPKGATDPGRPGEPLPG  
 ARGLTGRPGDAGPQKVGPTGAPGEDGRPGPPGPQGARGQPGVMGFPGPKGANGEPKAGEKGLPGAPGLRGLPGKDGETGAAGPP  
 GPAGPVGERGEQGAPCPGCFQGLPGPPGPPGESGKPGDQGVPGEAAPGLVGRGERGFPGERGSPGAQGLQGPRGLPGTPGTDGP  
 KGATGAPNGAQGPPGLQGMPGERGAAGIAGLKGDRGDVGEKGPEGAPGKDARGLTGPIGPPGPAGPNGEKGESGPPGSGAAG  
 ARGAPGERGEPEGAPCPAGFAGPPGADGQPGAKGEQGEQEPGQKGDAGAPGPQGPGAPGPQGPTGVTGPKGARGAQGPPGATGFPGAA  
 GRVGPPGPNGNPGPGPNGSAGKDGPKGVRGDAGPPGRAGDPGLQGPAGPPGEKGEPEGEDGPAGPDGPPGPQGLAGQRGIVGLPGQ  
 RGERGFPGLPGPSGEPKQGAPGSAGDRGPPGVPPGLTPAGEPGRGENPGADGLPGRDGAAGVKGDRGETGPVGAPGAPGAPG  
 APGPVGPTGKQGDRGETGAQGPMGPGSGPAGARGMPGPQGPRDKGETGEAGERGLKGHRGFTGLQGLPGPPGSDQGAAGPACPS  
 GPRGPPGPV

C-原肽

GPSGKDGSNGMPGPIGPPGPRGRS GEPGPAGPPGNGPPGPPGPP GTGIDMSAFAGLGQTE KGPDPIRYMRA DEA

C-原肽

AGGLRQHDVEVDATLKSLSNNQIESIRSPEGSKKNPARTCRDIKLCHPBWKSCDYWIDPNQGCTLDAIKVFCNMETGET  
 CVYPTPSSIPRKNWTSKTKDKHHWFATINGGFHFSYGDENLSPNTASIQMTFLRLLSTEGSQNVTYHCKNSIAYMDEETGNLK  
 KAILIQGSNDVEIRAEGNSRFTYSVLEDGCTKHTGKwGKTVIEYRLQKTSRLSIVDTAPMDIGGADQEFGVDIGPVCFL

FIG. 3 M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 M



FIG. 4



FIG. 5

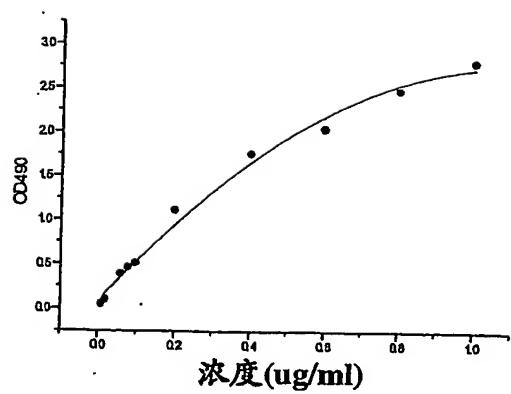


FIG. 6

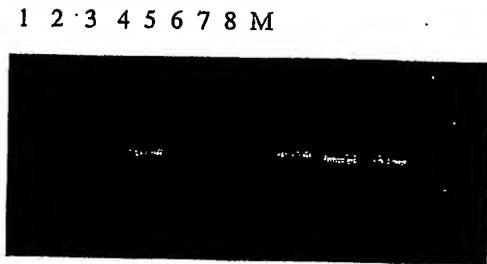


FIG. 7

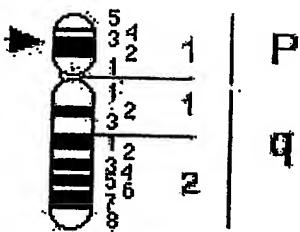


FIG. 8

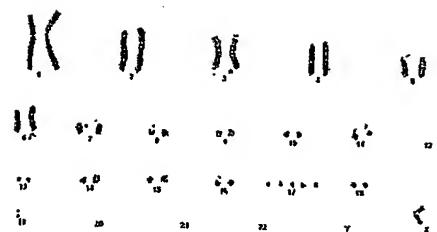


FIG. 9

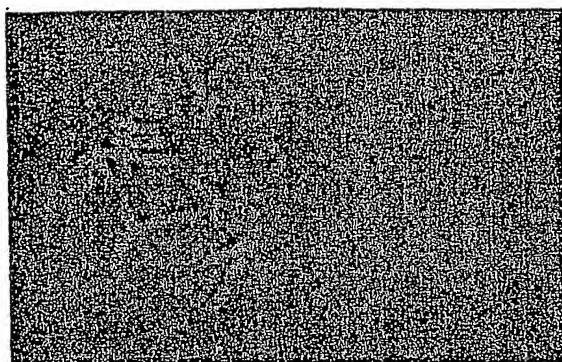


FIG. 10

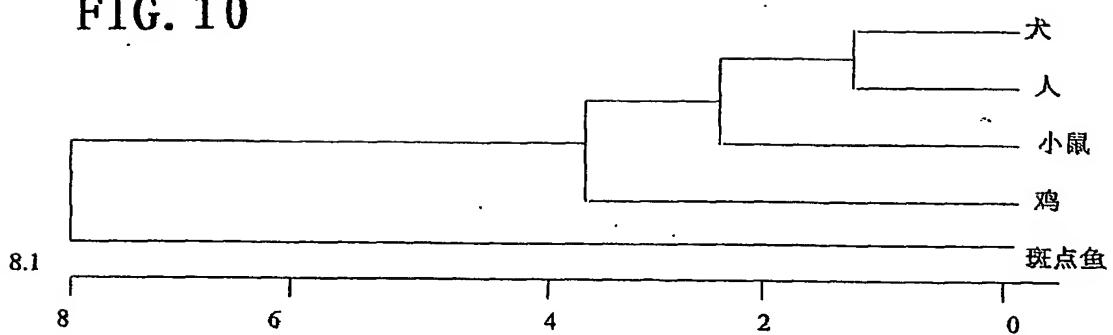


FIG. 11A

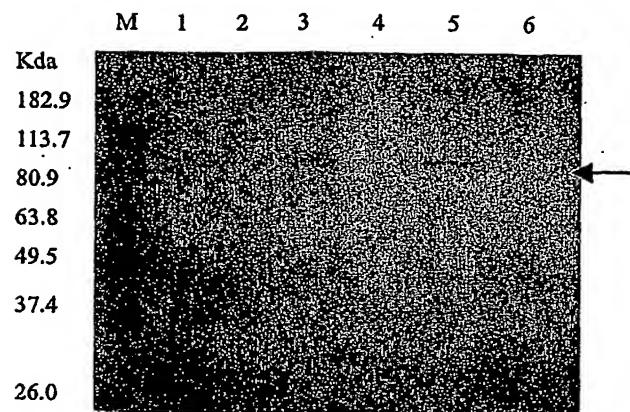


FIG. 11B

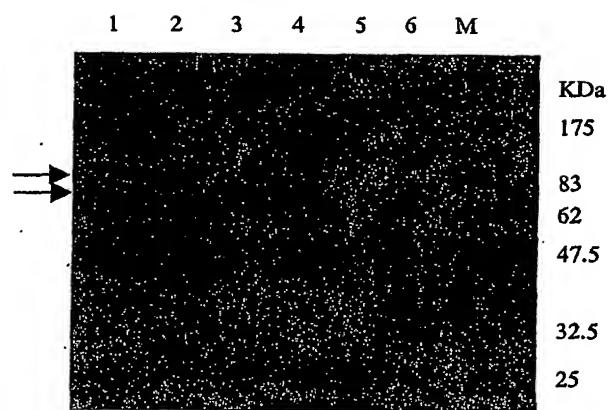
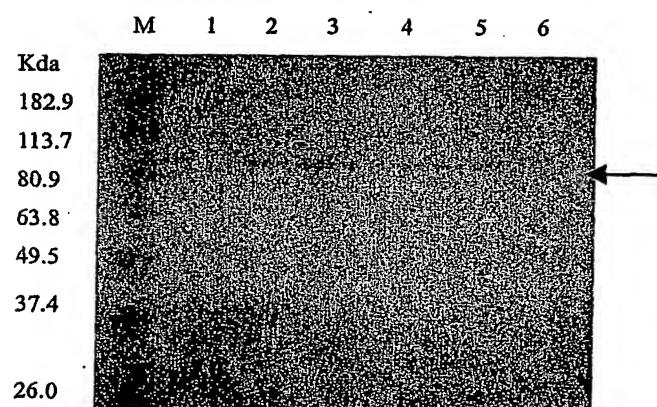


FIG. 12



## SEQUENCE LISTING

<110> 中国人民解放军军事医学科学院附属医院

<120> 一种编码鸡II型胶原蛋白的全长多核苷酸序列及其用途

<130> IEC030029PCT

<160> 3

<170> PatentIn version 3.1

<210> 1

<211> 5495

<212> DNA

<213> genomic DNA

<400> 1

ccaggcaagg atggcgcacg tgtaagtggg gcacggccat ggggtgggct ggcaaaggat 60

gctcacagag accacatcct catctctctc tctctccat agggtctgac gggtcccatt 120

ggtccccctg gccctgctgg ccccaacggt gagaaggtga gagcagcatc acagcacccc 180

acattacgccc ccatgggatg accccagtgc ctccacctct ccatcctttc ttttccaggg 240

tgaatccggc cctcctggtc catctgggtc tgccggtgcc cgtggtgccc ccgtaaagcac 300

aatgtctgca gcccctgggt gcccctaacc ttacccctaa acccccatac acccccttat 360

caacctcccc catctttcc cattaggggt agcgtggcga gcccggtgcc cccggtcctg 420

ctggatttgc tggcccccgg gtgagtgaaa cacccatcgac acacccacgt 480

tttcacccca catcctcacc ccactcatgg tggctgtgt tcccatcagg gcgcgcgtgg 540

acaacccgggt gccaaaggcg agcagggaga gcccgggcag aagggtgacg cgggcgtcc 600

tggtccccaa ggtccctccg gcgctctgg ccccccaggtaa caacacaaaa tggggcaaac 660

ccccaaatit gggacgtcac ggccccaaatg cagggcacact gcagctcccg ttggatttg 720

taacctgttt ttctctcattt cctagggtcc aaccgggtgtc actggtccca aaggagctcg 780

tggggctcag ggtccccctg tgagtaccgg ggggtgggct gcaggggtggg gaaggagcgg 840

ccgtggggct gagctgtgtc tgagccgtt ctccctccccc tcttcctct gactctgtga 900  
ttccctcccc agggagccac gggattcccc ggagctgccc gccgtgtggg accggccggc 960  
cctaattgtga gtctggggc gttctggat tgccccacc tggggtttg gcgctgcctc 1020  
cccgcgctgc gtgttggagg gggcactgtt tccctgcaca gacacgtggg gtttcctcc 1080  
ttggctctct gatgttggct tttggggca ttccaaatggt agagaaggac ttttctaagg 1140  
gcaagagctc cccaagaagc agcagtggga tgcgggtgat aaagatggaa tggctgcctc 1200  
tggttgcac caacgctgct ttccttcct ttaggtaac ccaggcccccc ccggacccccc 1260  
cggctctgct ggcaaagacg gccccaaaggg tggctgtggc gacgcccggcc ccccccggccg 1320  
tgcaggtgac cccggcctcc aaggccccgc cggccccccc ggcgagaagg gcgaacccgg 1380  
cgaggacggc cccgcggtaa ggattctggg ggtcttcctcc ctccgtgcac cccctggctg 1440  
cgtggtgccg ttgttcttag tctgatttcc ccctctgctg ccctgcaggg tcccgcggc 1500  
cccccccgcc cctcaaggct tggcaggaca gcgtggattt gtgggtctcc caggacagcg 1560  
tggtgagaga ggcttccccg gactgccggg gccatcggtg agtgggtcgc tctcatttgg 1620  
gtgcactgaa tcctatgggg tgcagagatg tggggccgc gatgctctgg agcccatctc 1680  
aggggtcgcc agccctttgg tgcagccggg ggacaccgtt tgcaagggtggg ttgggtttt 1740  
gcggagctcc ttttccccca ccaggagccg ctggtgcaag gcttaaagcc gggcaggaa 1800  
aaccatcagt gtttatttgt tgcagagggg tctggagcc ataaaaaaacg ggaaaggggc 1860  
agcgtgggg tctctccac tcatgcacct cttccatc tttcagggag aacctggaaa 1920  
gcaaggagcg cctggctctg cgggtgaccg aggtcccccc ggccccgtgg gccccctgg 1980  
gctgacaggt cctgctggag aaccggcg cgaggtaagc aaaacccac agcatcacag 2040  
cggcaccggg catcaccaac cccatggcac agtcagctc ccagagctcc .ccgggtgtctt 2100  
tttctccagc actgaaagga gactttgcac aaatccgtct ccacccgggt tgtaacatcc 2160

cctttcctc ctagggcaac cctggtgctg acggtcccc aggcaaggat ggcgcagctg 2220  
gcgtgaaggt gagcttgcca tgcgcctccc attggcactc gccatcccg tgccaaaagc 2280  
tgtgggttt tgcacagatc tgacctctct gttgtctgct cgcaaggta tcgtggtag 2340  
accggccctg tgggtgctcc cggtgctcct ggagccctg ggcgcggccgg ccctgttggt 2400  
cccactggaa aacaaggaga cagaggcgag acggtgagtg ctggcacaag ggtttagggt 2460  
ttagggtctc cttatggctg aaaatgtgca ggggttcccc tcaaggtttg ttccctgcac 2520  
cagtgctgag tgcatttaaa gatgctgtga ggcaccaaca gctgctgatt gtcaactgttg 2580  
cccgatctg gggtgcgagg catgggctg gctcagacac ccccgaaatc ccaaattcat 2640  
ggcttcgagg tggtgcttct ggtcgctggc accttctgat gtcctttt tctccctgca 2700  
gggtgcacaa gggccatgg gtccctctgg tcccgctgga gctcgaggaa tgccggtag 2760  
tggtgctgag tgcacatggca catcccacgt acagagcgtg gggtcctgctg tgccaggagg 2820  
gggtctgcca ccctgagccc gacacagccc tgtccccact ttagggtccc caaggacctc 2880  
gtggtgacaa aggtgagacg ggagaggctg gagagagagg gctgaaggc caccggcgct 2940  
tcaccggctc gcagggctcg cccggaccac ccgtaagttt gtttggggag cactgagccc 3000  
cccccccggt acgatgcggc tccttgggg tctctgtggc caccgaggct ctgtctggcc 3060  
caaagtgctg accgcagagc tgtgaccacc cggcttcct cctcagggcc cgtctggaga 3120  
ccaaggtgct gccggcccctc ctggccctc cggtcccaga gtaagtctg acggtggtgt 3180  
ttggggtggt ggaaggggaa ggagcagcag tggctccct gggcacctgc agcctctgtt 3240  
cgctccgtc tgctcatcag caccatgcc ttccctgccc tgaggcccg caatgcctc 3300  
acctcccggt tttggggctc tctcttaggg tcccccgtt cccgtcgcc cctctggcaa 3360  
agacggctct aacggcatgc cggcccat cggccctccc ggtccccgtg gacggagtgg 3420  
tgaacccggc cctgcgggtga gtcctggtaga ggggaggcag ggaatgggggt ccagctcgca 3480  
gagcagccca tcagcatcac ttcttctcc catagggtcc tcctggaaac cccggccctc 3540

ccggtcctcc tggccccccc ggcaccggca tcgacatgtc tgctttgct ggactgggtc 3600  
agacggagaa gggcccccgcac cccatccgct acatggggc agacgaggcg gccggagggc 3660  
tgccgcagca cgacgtggag gtggacgcca ccctcaaatac cctcaacaat cagattgaga 3720  
gcatccgcag ccccgagggc tccaagaaga accctgccag gacctgccgc gacatcaaac 3780  
tctgccatcc cgagtggaaag agcgtaaga gctccgcgtc cctctccgt cctccctct 3840  
tccccacagg agagcatccc cagcgtcctc gcaccgacct gcggtcaggt tggatgttag 3900  
gaaagattcc ttgtccaaaa gagctctggg cgctggcgtg ggctgcccgg ggaggtgggg 3960  
cagtcgctgt ccccataggt gttgggaac tttggagatg tggcacttgg gagcgtggct 4020  
tagtggggat gaggcagcag ttggaccaat ctgcagggtc ttctccagtc ttaatggctc 4080  
tgtgcttctg tcggtgtca tgggtgtat gggtggccat ttagacttgg cgatcttga 4140  
ggtctttcc gatcttaacg actcctagac ctcccaacc ccatgaacgc tggcttcct 4200  
ccccctgca ggagattact ggattgaccc gaaccaggc tgacacccgg acggcatcaa 4260  
agtattctgc aacatggaga caggcgagac ctgcgtctac ccgaccccca gcagcatccc 4320  
caggaagaac tgggtggacca gcaagacgaa agacaagaag cacgtctggg ttgcagagac 4380  
catcaacggc ggtttccacg tgggtgtccc cgggtgtcc ttggaaaggat cgatcccacc 4440  
tggatgtcc ttcttcgggt catgtggatg ggtttaatg aagttataga gggatgttct 4500  
gaaggtgtag gtttgggtca gttcagctcc acaaataaaa gggaaaggat gggatggagc 4560  
aactgagctc ctcggtttg ttggccag aaaaggttag gatgagggga ggcctcacgg 4620  
ccctacagcc ctttacggcc ctacagcagc gtttagaaaa aagttctgcc ccggagctgt 4680  
gttggcaca gaacagccct gtgtatggcgg agctcgggga gcattggac aacgctctca 4740  
gacattgggt ttgggtcagg tcctggtaa cgtgtatgtc agggggcaac cagcccatgg 4800  
gtggcctta aggacccttc caagccaaacc attccatggt tctgtatct gtaaggaccc 4860

|  |      |
|--|------|
| ttccaaatcca aaccactctg attttttct cagccatttg ggaacctgaa gtacggaagt  | 4920 |
| cctccccaaa agtcctgag agtaagggtgg tcataatgcc cgcaaggcctt aactcctcac | 4980 |
| ctctccctc cagttcagct acggcgatga gaacctgtcc cccaaacaccg ccagcatcca  | 5040 |
| gatgacccctc ctgcgcctcc tgtccaccga gggctcccag aacgtcacct accactgcaa | 5100 |
| gaacagcatc gcctacatgg acgaggagac gggcaacctg aagaaagcca tcctcatcca  | 5160 |
| gggatccaac gacgtggaga tcagagccga gggcaacagc aggtcacct acagcgtctt   | 5220 |
| ggaggacggc tgcacggtag gttgctgggc gcctgcaaag gaaaggtgca gatggggagg  | 5280 |
| gggaggctga ggctgggggg atgaggccgg agcagctgac agcatccctg ccctccctcc  | 5340 |
| ctccccagaa acacactggc aaatggggca agacggtgat cgagtaccgg tcgcagaaga  | 5400 |
| cctcgccct gcccattgta gatattgcac ctatggacat tggcggagcc gatcaggagt   | 5460 |
| ttggcgtgga tattggccca gtctgcttct tgtaa                             | 5495 |

&lt;210&gt; 2

&lt;211&gt; 4793

&lt;212&gt; DNA

&lt;213&gt; cDNA

&lt;400&gt; 2

|  |     |
|--|-----|
| atgcacggcc gccgcccccc ccgctccgcc gctctccctcc tcctccctct ccttctcacg | 60  |
| gccgccccaa ccgcgcagga ccgcgcaccc cgacaacctg gccccaaaggg acagaaggga | 120 |
| gaacccggag atattaaaga tggtagga ccccgagggc ctccaggacc acagggccca    | 180 |
| gcaggagagc agggacacgc aggggaccgt ggcgagaagg gggagaagg tgctccctggc  | 240 |
| ccccgtggga gggatggaga acccggcacc cctggaaacc caggcccccc cggtcccccc  | 300 |
| ggacccctcg gcccccccg acttggtgga aactttgcgg cgcagatggc gggccggcttc  | 360 |
| gatgagaagg cgggtggagc gcagatgggt gtcatgcagg gaccatggg ccctatggga   | 420 |
| cccccgccccc cccctggccc cactggcgca cctggtcccc agggattca aggcaacccc  | 480 |

ggtagccccg gcaaacccgg cgctgtggc ccgatgggt cccggggacc tccgggacca 540  
cctggaaac ccggtgacga tggtagaca ggcaaaccg gcaaatctgg tgaacgtggc 600  
ccccccggcc cccagggcgc tcgtggcttc cctggactc ctggtctccc cggagtgaag 660  
ggccaccgag gctaccccggttttggatggt gc当地aggag aggccggggc tcctggagcc 720  
aagggtgaat ctggttcacc gggtgagaac ggctcccccggcccccattggg accccgtggg 780  
ctgccccggag agcgaggacg tccggcccc tccggcgccg ccggtgctcg tggcaatgac 840  
ggtctccctg gccctgctgg accccctggg cccgtcgccctc ctggggagc cccggcttc 900  
cccgagccccccgggtcaaaa gggtgaagcc ggccccactg gtgcacgggg tcccgagggt 960  
gcccaaggac cccgcggcga atccggcacc cccggctctc ccggcccccgc tggcgcaccc 1020  
ggtaaccaggag ggactgtatgg catcccggtt gccaagggtt cggcgggtgc cccgggcatt 1080  
gcaggcgctc caggattcccc cggccacgcg ggccccccgg gaccccaagg tgccaccgg 1140  
ccactgggac ccaaaggaca gacggcgaa cccggcatcg caggcttcaa gggcgagcaa 1200  
ggaccgaagg gcgagacggg cccgcagga ccccaagggtt ccccccggcc ggctgggttag 1260  
gaaggcaaga gaggagctcg tggtagacct ggtgccggcc gccctgtggg ccccccgg 1320  
gaaagggcgctcctggcaa ccgtggattc cccggcagg acgggctggc cggacccaaag 1380  
gggtctccag gtgaacgcgg cccgcgtgg tctggccgtc ccaaagggtgc caccgggtac 1440  
cccgacgtc ccggagagcc cggcgtggcc ggagcgaggg gtctcaccgg ccggccggc 1500  
gatgcgggac ctcaaggcaa agtcggccaa actggtgctc ctggcgagga tggccggccc 1560  
ggccccccgg gacctcaggg tgctcggtt cggctgggtt tggatgggtt cccgggtccc 1620  
aaaggcgcta atggtgagcc tggaaaagct ggagagaaag gactgccccgg cgccccagg 1680  
ctgcggggtc tgcctggcaa ggatggggag acgggagctg ccggccccc tggacccgct 1740  
ggtcctgtgg gtgagagagg agagcaagga gccccggc cttccggctt ccaggactg 1800  
cccgaccac cagggtcccccc tggggagagc ggcaaaccg gagaccagggtt tggatgg 1860

gaagccggtg cccccggtgtct tggtggtccc agaggtgaac gtggattccc cggtgaacgc 1920  
ggctctcccg gtgccccagg gctgcagggt ccccggtggc tccccggAAC gcccggcact 1980  
gacggaccCA agggtgcaac cggTccAGCC ggccccAAACg gtgcccAGGG tcccccAGGG 2040  
ctgcaggGAA tgcccggta gagaggAGCA gctggcatcg ctggcctcaa gggtgaccGG 2100  
ggagatgttG gtgagaaaAGG acctgaggGA gctccaggCA aggatggcgc acgtggtctG 2160  
acgggtcccc ttggtccccC tggccctgct ggccccAAACg gtgagaAGGG tgaatccGGC 2220  
cctcctggTC catctggTGC tgccggtgCC cgtggtgccc ccggtgagCG tggcgagccc 2280  
ggtgccccCG gtcctgctGG atttgctggc cccccggcG ccgatggaca acccggtGCC 2340  
aaaggcgagC aggagAGGC cgggcagaAG ggtgacgCgg ggcctcctGG tccccAAGGT 2400  
ccctccggCG ctccctggccc ccagggcccA accgggtGTCA ctggtccccAA aggagctcGT 2460  
ggggctcagg gtccccctGG agccacggGA ttccccggAG ctggcggccc tgtgggaccG 2520  
cccgccccCTA atggtaACCC aggccccccc ggaccccccTG gctctgctGG caaggacGGC 2580  
cccaagggtG ttctgtggcGA cgccggcccc cccggccGTG caggtgacCC cggcctccAA 2640  
ggccccggCG gccccccCG CGAGAAGGGC gaacccggcG aggacggccc cgggggtccc 2700  
gacggcccccc cggccctcA aggcttggcA ggacagcgtG gtattgtggG tctcccAGGA 2760  
cagcgtggTG agagaggcTT ccccgactG cggggccat cgggagaACC tggaaAGCAA 2820  
ggagcgcctG gctctgcggG tgaccgagGT ccccccggcC ccgtggggcc ccctggcTG 2880  
acgggtcctG ctggagaACC cgggcgcgAG ggcaaccctG gtgcgtacGG tctcccAGGC 2940  
aggatggcG cagctggcGT gaagggtGAT cgtggtgAGA ccggccctGT gggtgccccC 3000  
ggtgctcctG gagccccctGG cgccccggc CCTgtggTC ccactggAAA acaaggagAC 3060  
agaggcgaga cgggtgcACA agggcccatG ggtccctcTG gtcccgtGG agctcgagGA 3120  
atgccgggtc cccaaggacc tcgtggtgac aaaggtgaga cgggagaggc tggagagaga 3180



|   |      |
|---|------|
| atcaatctgg caggtgtgac ggcccccctc cccacaaaagg gatctggcaa acgcaggtat  | 4620 |
| cgcgaatccc ctccccctccc cgtgtatcac cagcaggagt gctaatgtat cataacaacag | 4680 |
| aaatggtgct attcttgtaa aacaagtctg tatttttaa catcagttga tataaaaaca    | 4740 |
| acaaaaaaaaaa aaacttttgg tggaaagtaa aaaaaacaaa aaaaaaaaaa aaa        | 4793 |

<210> 3  
<211> 1420  
<212> PRT  
<213> chicken

<400> 3

|   |   |    |    |
|---|---|----|----|
| Met His Gly Arg Arg Pro Pro Arg Ser Ala Ala Leu Leu Leu Leu |   |    |    |
| 1   | 5 | 10 | 15 |

|   |    |    |
|---|----|----|
| Leu Leu Leu Thr Ala Ala Ala Ala Gln Asp Arg Asp Leu Arg Gln |    |    |
| 20  | 25 | 30 |

|   |    |    |
|---|----|----|
| Pro Gly Pro Lys Gly Gln Lys Gly Glu Pro Gly Asp Ile Lys Asp Val |    |    |
| 35  | 40 | 45 |

|   |    |    |
|---|----|----|
| Val Gly Pro Arg Gly Pro Pro Gly Pro Gln Gly Pro Ala Gly Glu Gln |    |    |
| 50  | 55 | 60 |

|   |    |    |    |
|---|----|----|----|
| Gly Gln Arg Gly Asp Arg Gly Glu Lys Gly Glu Lys Gly Ala Pro Gly |    |    |    |
| 65  | 70 | 75 | 80 |

|   |    |    |
|---|----|----|
| Pro Arg Gly Arg Asp Gly Glu Pro Gly Thr Pro Gly Asn Pro Gly Pro |    |    |
| 85  | 90 | 95 |

|   |     |     |
|---|-----|-----|
| Pro Gly Pro Pro Gly Pro Pro Gly Leu Gly Gly Asn Phe |     |     |
| 100   | 105 | 110 |

Ala Ala Gln Met Ala Gly Gly Phe Asp Glu Lys Ala Gly Gly Ala Gln  
115 120 125

Met Gly Val Met Gln Gly Pro Met Gly Pro Met Gly Pro Arg Gly Pro  
130 135 140

Pro Gly Pro Thr Gly Ala Pro Gly Pro Gln Gly Phe Gln Gly Asn Pro  
145 150 155 160

Gly Glu Pro Gly Glu Pro Gly Ala Ala Gly Pro Met Gly Pro Arg Gly  
165 170 175

Pro Pro Gly Pro Pro Gly Lys Pro Gly Asp Asp Gly Glu Thr Gly Lys  
180 185 190

Pro Gly Lys Ser Gly Glu Arg Gly Pro Pro Gly Pro Gln Gly Ala Arg  
195 200 205

Gly Phe Pro Gly Thr Pro Gly Leu Pro Gly Val Lys Gly His Arg Gly  
210 215 220

Tyr Pro Gly Leu Asp Gly Ala Lys Gly Glu Ala Gly Ala Pro Gly Ala  
225 230 235 240

Lys Gly Glu Ser Gly Ser Pro Gly Glu Asn Gly Ser Pro Gly Pro Met  
245 250 255

Gly Pro Arg Gly Leu Pro Gly Glu Arg Gly Arg Pro Gly Pro Ser Gly  
260 265 270

Ala Ala Gly Ala Arg Gly Asn Asp Gly Leu Pro Gly Pro Ala Gly Pro  
275 280 285

Pro Gly Pro Val Gly Pro Ala Gly Ala Pro Gly Phe Pro Gly Ala Pro

290                    295                    300

Gly Ser Lys Gly Glu Ala Gly Pro Thr Gly Ala Arg Gly Pro Glu Gly  
305                    310                    315                    320

Ala Gln Gly Pro Arg Gly Glu Ser Gly Thr Pro Gly Ser Pro Gly Pro  
325                    330                    335

Ala Gly Ala Pro Gly Asn Pro Gly Thr Asp Gly Ile Pro Gly Ala Lys  
340                    345                    350

Gly Ser Ala Gly Ala Pro Gly Ile Ala Gly Ala Pro Gly Phe Pro Gly  
355                    360                    365

Pro Arg Gly Pro Pro Gly Pro Gln Gly Ala Thr Gly Pro Leu Gly Pro  
370                    375                    380

Lys Gly Gln Thr Gly Glu Pro Gly Ile Ala Gly Phe Lys Gly Glu Gln  
385                    390                    395                    400

Gly Pro Lys Gly Glu Thr Gly Pro Ala Gly Pro Gln Gly Ala Pro Gly  
405                    410                    415

Pro Ala Gly Glu Glu Gly Lys Arg Gly Ala Arg Gly Glu Pro Gly Ala  
420                    425                    430

Ala Gly Pro Val Gly Pro Pro Gly Glu Arg Gly Ala Pro Gly Asn Arg  
435                    440                    445

Gly Phe Pro Gly Gln Asp Gly Leu Ala Gly Pro Lys Gly Ala Pro Gly  
450                    455                    460

Glu Arg Gly Pro Ala Gly Leu Ala Gly Pro Lys Gly Ala Thr Gly Asp  
465                    470                    475                    480

Pro Gly Arg Pro Gly Glu Pro Gly Leu Pro Gly Ala Arg Gly Leu Thr  
485 490 495

Gly Arg Pro Gly Asp Ala Gly Pro Gln Gly Lys Val Gly Pro Thr Gly  
500 505 510

Ala Pro Gly Glu Asp Gly Arg Pro Gly Pro Pro Gly Pro Gln Gly Ala  
515 520 525

Arg Gly Gln Pro Gly Val Met Gly Phe Pro Gly Pro Lys Gly Ala Asn  
530 535 540

Gly Glu Pro Gly Lys Ala Gly Glu Lys Gly Leu Pro Gly Ala Pro Gly  
545 550 555 560

Leu Arg Gly Leu Pro Gly Lys Asp Gly Glu Thr Gly Ala Ala Gly Pro  
565 570 575

Pro Gly Pro Ala Gly Pro Val Gly Glu Arg Gly Glu Gln Gly Ala Pro  
580 585 590

Gly Pro Ser Gly Phe Gln Gly Leu Pro Gly Pro Pro Gly Pro Pro Gly  
595 600 605

Glu Ser Gly Lys Pro Gly Asp Gln Gly Val Pro Gly Glu Ala Gly Ala  
610 615 620

Pro Gly Leu Val Gly Pro Arg Gly Glu Arg Gly Phe Pro Gly Glu Arg  
625 630 635 640

Gly Ser Pro Gly Ala Gln Gly Leu Gln Gly Pro Arg Gly Leu Pro Gly  
645 650 655

Thr Pro Gly Thr Asp Gly Pro Lys Gly Ala Thr Gly Pro Ala Gly Pro  
660                    665                    670

Asn Gly Ala Gln Gly Pro Pro Gly Leu Gln Gly Met Pro Gly Glu Arg  
675                    680                    685

Gly Ala Ala Gly Ile Ala Gly Leu Lys Gly Asp Arg Gly Asp Val Gly  
690                    695                    700

Glu Lys Gly Pro Glu Gly Ala Pro Gly Lys Asp Gly Ala Arg Gly Leu  
705                    710                    715                    720

Thr Gly Pro Ile Gly Pro Pro Gly Pro Ala Gly Pro Asn Gly Glu Lys  
725                    730                    735

Gly Glu Ser Gly Pro Pro Gly Pro Ser Gly Ala Ala Gly Ala Arg Gly  
740                    745                    750

Ala Pro Gly Glu Arg Gly Glu Pro Gly Ala Pro Gly Pro Ala Gly Phe  
755                    760                    765

Ala Gly Pro Pro Gly Ala Asp Gly Gln Pro Gly Ala Lys Gly Glu Gln  
770                    775                    780

Gly Glu Pro Gly Gln Lys Gly Asp Ala Gly Ala Pro Gly Pro Gln Gly  
785                    790                    795                    800

Pro Ser Gly Ala Pro Gly Pro Gln Gly Pro Thr Gly Val Thr Gly Pro  
805                    810                    815

Lys Gly Ala Arg Gly Ala Gln Gly Pro Pro Gly Ala Thr Gly Phe Pro  
820                    825                    830

Gly Ala Ala Gly Arg Val Gly Pro Pro Gly Pro Asn Gly Asn Pro Gly  
835                    840                    845

Pro Pro Gly Pro Pro Gly Ser Ala Gly Lys Asp Gly Pro Lys Gly Val  
850                    855                    860

Arg Gly Asp Ala Gly Pro Pro Gly Arg Ala Gly Asp Pro Gly Leu Gln  
865                    870                    875                    880

Gly Pro Ala Gly Pro Pro Gly Glu Lys Gly Glu Pro Gly Glu Asp Gly  
885                    890                    895

Pro Ala Gly Pro Asp Gly Pro Pro Gly Pro Gln Gly Leu Ala Gly Gln  
900                    905                    910

Arg Gly Ile Val Gly Leu Pro Gly Gln Arg Gly Glu Arg Gly Phe Pro  
915                    920                    925

Gly Leu Pro Gly Pro Ser Gly Glu Pro Gly Lys Gln Gly Ala Pro Gly  
930                    935                    940

Ser Ala Gly Asp Arg Gly Pro Pro Gly Pro Val Gly Pro Pro Gly Leu  
945                    950                    955                    960

Thr Gly Pro Ala Gly Glu Pro Gly Arg Glu Gly Asn Pro Gly Ala Asp  
965                    970                    975

Gly Leu Pro Gly Arg Asp Gly Ala Ala Gly Val Lys Gly Asp Arg Gly  
980                    985                    990

Glu Thr Gly Pro Val Gly Ala Pro Gly Ala Pro Gly Ala Pro Gly Ala  
995                    1000                    1005

Pro Gly Pro Val Gly Pro Thr Gly Lys Gln Gly Asp Arg Gly Glu

1010                    1015                    1020

Thr Gly Ala Gln Gly Pro Met Gly Pro Ser Gly Pro Ala Gly Ala  
1025                    1030                    1035

Arg Gly Met Pro Gly Pro Gln Gly Pro Arg Gly Asp Lys Gly Glu  
1040                    1045                    1050

Thr Gly Glu Ala Gly Glu Arg Gly Leu Lys Gly His Arg Gly Phe  
1055                    1060                    1065

Thr Gly Leu Gln Gly Leu Pro Gly Pro Pro Gly Pro Ser Gly Asp  
1070                    1075                    1080

Gln Gly Ala Ala Gly Pro Ala Gly Pro Ser Gly Pro Arg Gly Pro  
1085                    1090                    1095

Pro Gly Pro Val Gly Pro Ser Gly Lys Asp Gly Ser Asn Gly Met  
1100                    1105                    1110

Pro Gly Pro Ile Gly Pro Pro Gly Pro Arg Gly Arg Ser Gly Glu  
1115                    1120                    1125

Pro Gly Pro Ala Gly Pro Pro Gly Asn Pro Gly Pro Pro Gly Pro  
1130                    1135                    1140

Pro Gly Pro Pro Gly Thr Gly Ile Asp Met Ser Ala Phe Ala Gly  
1145                    1150                    1155

Leu Gly Gln Thr Glu Lys Gly Pro Asp Pro Ile Arg Tyr Met Arg  
1160                    1165                    1170

Ala Asp Glu Ala Ala Gly Gly Leu Arg Gln His Asp Val Glu Val  
1175                    1180                    1185

Asp Ala Thr Leu Lys Ser Leu Asn Asn Gln Ile Glu Ser Ile Arg  
1190 1195 1200

Ser Pro Glu Gly Ser Lys Lys Asn Pro Ala Arg Thr Cys Arg Asp  
1205 1210 1215

Ile Lys Leu Cys His Pro Glu Trp Lys Ser Gly Asp Tyr Trp Ile  
1220 1225 1230

Asp Pro Asn Gln Gly Cys Thr Leu Asp Ala Ile Lys Val Phe Cys  
1235 1240 1245

Asn Met Glu Thr Gly Glu Thr Cys Val Tyr Pro Thr Pro Ser Ser  
1250 1255 1260

Ile Pro Arg Lys Asn Trp Trp Thr Ser Lys Thr Lys Asp Lys Lys  
1265 1270 1275

His Val Trp Phe Ala Glu Thr Ile Asn Gly Gly Phe His Phe Ser  
1280 1285 1290

Tyr Gly Asp Glu Asn Leu Ser Pro Asn Thr Ala Ser Ile Gln Met  
1295 1300 1305

Thr Phe Leu Arg Leu Leu Ser Thr Glu Gly Ser Gln Asn Val Thr  
1310 1315 1320

Tyr His Cys Lys Asn Ser Ile Ala Tyr Met Asp Glu Glu Thr Gly  
1325 1330 1335

Asn Leu Lys Lys Ala Ile Leu Ile Gln Gly Ser Asn Asp Val Glu  
1340 1345 1350

Ile Arg Ala Glu Gly Asn Ser Arg Phe Thr Tyr Ser Val Leu Glu  
1355 1360 1365

Asp Gly Cys Thr Lys His Thr Gly Lys Trp Gly Lys Thr Val Ile  
1370 1375 1380

Glu Tyr Arg Leu Gln Lys Thr Ser Arg Leu Ser Ile Val Asp Thr  
1385 1390 1395

Ala Pro Met Asp Ile Gly Gly Ala Asp Gln Glu Phe Gly Val Asp  
1400 1405 1410

Ile Gly Pro Val Cys Phe Leu  
1415 1420

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN03/00967

## A. CLASSIFICATION OF SUBJECT MATTER

IPC<sup>7</sup> C07H21/04 C07K14/465 C07K14/78 C12N15/63 A61K38/17 A61K38/39 A61P19/02  
 According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC<sup>7</sup> C07H21/04 C07K14/465 C07K14/78 C12N15/63 A61K38/17 A61K38/39 A61P19/02

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT WPI EPODOC NCBI

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                | Relevant to claim No. |
|-----------|---|-----------------------|
| A         | WO02058484A2 01 Aug 2002, Refer to the Abstract   | 1-10                  |
| A         | US6323319B1 27 Nov 2001, Refer to the Abstract  | 1-10                  |
| A         | NCBI Register No: L00061.1 GI:211338 (Ninomiya, Y., et al)<br>25 April 2002, Refer to the SEQ ID. | 1-10                  |

Further documents are listed in the continuation of Box C.  See patent family annex.

|  |  |
|--|--|
| * Special categories of cited documents: |  |
| "A"                                      | document defining the general state of the art which is not considered to be of particular relevance   |
| "E"                                      | earlier application or patent but published on or after the international filing date  |
| "L"                                      | document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)   |
| "O"                                      | document referring to an oral disclosure, use, exhibition or other means   |
| "P"                                      | document published prior to the international filing date but later than the priority date claimed   |
| "T"                                      | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  |
| "X"                                      | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| "Y"                                      | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| "&"                                      | document member of the same patent family  |

Date of the actual completion of the international search  
25 Mar 2004(25.03.04)

Date of mailing of the international search report  
15 · APR 2004 (15 · 04 · 2004)

Name and mailing address of the ISA/CN  
6 Xitucheng Rd., Jimen Bridge, Haidian District,  
100088 Beijing, China  
Facsimile No. 86-10-62019451

Authorized officer

Guo Xiaoyong  
Telephone No. 8610-62085057



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/CN03/00967

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos:11  
because they relate to subject matter not required to be searched by this Authority, namely:  
Claim 11 is directed to a method of treatment of the human/animal body.
2.  Claims Nos:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.  Claims Nos:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently ,this international search report is restricted to the invention first mentioned in the claims;it is covered by claims Nos.:

#### Remark on protest

The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

PCT / CN03/00967

WO02058484A2 Publishing Date: 01 Aug 2002 Patent family members:  
US2002137688A1 Publishing Date: 26 Sept 2002

## 国际检索报告

国际申请号

PCT/CN03/00967

## A. 主题的分类

IPC<sup>7</sup> C07H21/04 C07K14/465 C07K14/78 C12N15/63 A61K38/17 A61K38/39 A61P19/02

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

## B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

IPC<sup>7</sup> C07H21/04 C07K14/465 C07K14/78 C12N15/63 A61K38/17 A61K38/39 A61P19/02

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

CNPAT WPI EPODOC NCBI

## C. 相关文件

| 类 型* | 引用文件, 必要时, 指明相关段落   | 相关的权利要求编号 |
|------|---|-----------|
| A    | WO02058484A2 01.8 月 2002, 参见摘要                                    | 1-10      |
| A    | US6323319B1 27.11 月 2001, 参见摘要                                    | 1-10      |
| A    | NCBI 登记号 L00061.1 GI:211338 (Ninomiya,Y.,等人)<br>25.4 月 2002, 参见序列 | 1-10      |

 其余文件在 C 栏的续页中列出。 见同族专利附件。

\* 引用文件的专用类型:

“A” 明确叙述了被认为不是特别相关的一般现有技术的文件

“E” 在国际申请日的当天或之后公布的在先的申请或专利

“L” 可能引起对优先权要求的怀疑的文件, 为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理

“X” 特别相关的文件, 仅仅考虑该文件, 权利要求所记载的发明就不能认为是新颖的或不能认为是有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 权利要求记载的发明不具有创造性

“&amp;” 同族专利成员的文件

国际检索实际完成的日期

25.3 月 2004(25.3.04)

国际检索报告邮寄日期

15 : 4月 2004 (15 · 04 · 2004)

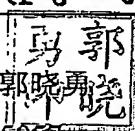
国际检索单位名称和邮寄地址

ISA/CN

中国北京市海淀区西土城路 6 号(100088)

传真号: 86-10-62019451

受权官员


  
郭晓勇

电话号码: 86-10-62085057

**第I栏 关于某些权利要求不能作为检索主题的意见(接第 1 页第 1 项)**

按条约 17(2)(a)对某些权利要求未作国际检索报告的理由如下:

1.  权利要求 (编号) : 11

因为它们涉及到不要求本国际检索单位检索的主题, 即:  
其涉及疾病的治疗和诊断方法。

2.  权利要求 (编号) :

因为它们涉及到国际申请中不符合规定的要求的部分, 以至于不能进行任何有意义的国际检索,  
具体地说:

3.  权利要求 (编号) :

因为它们是从属权利要求, 并且没有按照细则 6.4(a)第 2 句和第 3 句的要求撰写。

**第II栏 关于缺乏发明单一性时的意见(接第 1 页第 2 项)**

本国际检索单位在该国际申请中发现多项发明, 即:

1.  由于申请人按时缴纳了所要求缴纳的全部附加检索费, 本国际检索报告针对全部可作检索的权利要求。

2.  由于无需付出有理由要求附加费的劳动即能对全部可检索的权利要求都进行检索, 本国际检索单位未通知缴纳任何附加费。

3.  由于申请人仅按时缴纳了部分所要求缴纳的附加检索费, 本国际检索报告仅涉及已缴费的那些权利要求。具体地说, 是权利要求 (编号) :

4.  申请人未按时缴纳所要求的附加检索费。因此, 本国际检索报告仅涉及权利要求中首先提到的发明;  
包含该发明的权利要求是 (编号) :

关于异议的说明:  申请人的异议书随附加检索费同时提交。  
 支付附加检索费时未提交异议书。

国际检索报告  
关于同族专利成员的情报

国际申请号  
PCT/CN03/00967

| 检索报告中引用的<br>专利文件 | 公布日期        | 同族专利成员         | 公布日期        |
|------------------|-------------|----------------|-------------|
| WO02058484A2     | 01.8 月 2002 | US2002137688A1 | 26.9 月 2002 |